

IFUP-TH/20124

SELECTED TOPICS IN QUANTUM MECHANICS

Pietro Menotti

Dipartimento di Fisica, Università di Pisa, Italy

September 2024

INDEX

Foreword

Chapter I:	Remarks on the black-body radiation	7
	1. Introduction	
	2. Wien's law and the number of photons	
	3. Energy fluctuations in the black body	
	References	
Chapter II:	The superposition principle	15
	1. Observables, spectrum, eigenstates and probability	
	2. The superposition principle	
	3. Standard expansion of a vector	
	4. Observable with bounded but otherwise arbitrary spectrum	
	5. Continuum spectrum in the Dirac formalism	
	References	
Chapter III:	The correspondence principle	25
	References	
Chapter IV:	The time evolution	31
	1. Introduction	
	2. Composition properties of U	
	3. Construction of the solution to the time evolution equation	
	4. Expansion in T -ordered integrals	
	References	
Chapter V:	Properties of the hamiltonian operator	37
	1. Introduction	
	2. Some properties of the spectrum	
	3. Lower boundedness of the spectrum	
	References	

Chapter VI: Description of composite systems	43
Chapter VII: The density operator	47
1. Introduction	
2. The time evolution of the density operator	
References	
Chapter VIII: Measure in quantum mechanics	53
References	
Chapter IX: Symmetry transformations	59
1. Space translations	
2. Symmetry transformations	
3. Rotations	
4. The euclidean group	
References	
Chapter X: The spin	71
1. Realization of half-integer angular momenta	
Chapter XI: Identical particles	77
1. Introduction	
2. Two particle system	
3. Three particle system	
References	
Chapter XII: The formal theory of scattering	91
1. Introduction	
2. The Lippmann-Schwinger equation	
3. The adiabatic theorem	
4. “In” and “out” states	

5. The Møller wave function operators and the S -matrix	
6. The unitarity of the S -matrix and the optical theorem	
7. Transition probabilities per unit time	
References	
Chapter XIII: Scattering by a central potential	107
1. Introduction	
2. The Jost functions	
3. Physical interpretation of the singularities	
4. Levinson's theorem	
5. Construction of special potentials	
6. Discussion of the singularities of the S -matrix	
References	
Chapter XIV: The functional formulation of quantum mechanics	129
1. Introduction	
2. The path integral	
3. Proof of Trotter formula	
References	
Chapter XV: The violation of Bell's inequalities	141
1. Introduction	
2. Bell's original argument	
3. The Clauser-Holt-Horn-Shimony inequality	
References	

Foreword

The books adopted in the course on Quantum Mechanics are

P.A.M. Dirac *The Principles of Quantum Mechanics*, Oxford at the Clarendon Press (1958)

L.D. Landau and E.M. Lifshitz *Quantum Mechanics*, Pergamon Press (1962)

These notes are meant as complements to the above two books

Reference to other books are found at the end of each chapter of these notes.

Chapter 1

Remarks on the black-body radiation

1.1 Introduction

The energy density of the electromagnetic field is given by

$$u = \frac{\epsilon_0}{2}(E^2 + c^2 B^2) \text{ (MKS)}; \quad u = \frac{1}{8\pi}(E^2 + B^2) \text{ (Gauss)} .$$

It is a quadratic form in the fields E_j, B_j . As such the time average of the energy possesses a spectral decomposition.

Given $f(t)$, $-L/2 < t < L/2$ let us write

$$f(t) = \frac{1}{\sqrt{L}} \sum_n e^{i\omega_n t} \tilde{f}(\omega_n); \quad \tilde{f}(\omega_n) = \frac{1}{\sqrt{L}} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(t) e^{-i\omega_n t} dt$$

$\omega_n = 2\pi n/L$. $\tilde{f}(-\omega) = \tilde{f}^*(\omega)$ due to the reality of f . The time average value is

$$\frac{1}{L} \int_{-L/2}^{L/2} |f(t)|^2 dt = \frac{1}{L} \sum_n |\tilde{f}(\omega_n)|^2 \rightarrow \frac{1}{2\pi} \int |\tilde{f}(\omega)|^2 d\omega = \frac{1}{\pi} \int_0^\infty |\tilde{f}(\omega)|^2 d\omega$$

where the arrow denotes the transition to the continuum frequency case.

Thus we can write

$$u(T) = \int_0^\infty u_\nu(\nu, T) d\nu$$

where T is the temperature.

Experimentally equilibrium in a black body is reached in very short times, of the order of l/c being l the linear dimension of the black body.

Kirchhoff's theorems assure the universality of $u_\nu(\nu, T)$, i.e. the independence from the material of the walls, volume, shape of the cavity and from the position of the detected radiation.

The proof relies on the second principle of thermodynamics.

The Stefan-Boltzmann law states

$$u(T) = K_1 T^4 \quad (1.1)$$

The derivation is based on

- 1) The second principle of thermodynamics $Q/T = dS$;
- 2) The electromagnetic nature of the radiation: in fact from the expression of the momentum density

$$P = \epsilon_0(E \wedge B) \text{ (MKS)}; \quad P = \frac{1}{4\pi c}(E \wedge B) \text{ (Gauss)}$$

we have for a plane wave

$$|P| = \frac{u}{c}$$

and for an isotropic radiation it follows for the pressure $p = u/3$.

Exploiting the exactness of $dS = \frac{1}{T}(dU + pdV)$ with $U = u(T)V$, $p = \frac{u(T)}{3}$ we have

$$\frac{du}{u} = 4 \frac{dT}{T}$$

from which we have (1.1). Moreover by integration one obtains for the entropy

$$S = \frac{4}{3} K_1 V T^3 + \text{const.} \quad (1.2)$$

The radiance of the black body is given by

$$\text{power} = \sigma T^4 \times \text{area}$$

where σ is the Stefan constant related to K_1 by the kinematical relation $\sigma = cK_1/4$. Experimentally $\sigma = 5.67 \cdot 10^{-8} \text{ J/m}^2 \text{ s K}^4$.

Example: The temperature of the sun is $T = 6000\text{ K}$. The apparent radius of the sun is $1/4$ of degree $= 0.0043$. It follows that the power of the sun radiation on the earth is

$$w = \sigma (0.0043)^2 (6000)^4\text{ K}^4 = 1358\text{ watt/m}^2 .$$

Wien's law states that

$$u_\nu(\nu, T) = \nu^3 f\left(\frac{\nu}{T}\right) \quad (1.3)$$

and it is derived from [1]

- 1) Second principle of thermodynamics;
- 2) Doppler effect.

First part: If one compresses slowly the black radiation in a vessel with reflecting walls the radiation stays black; only its temperature changes according to the constancy of the entropy (1.2).

Second part: Render this compatible with the Doppler effect

$$\nu' = \nu \left(1 + 2\frac{v}{c} \cos \theta\right); \quad (v/c \ll 1) .$$

where v is the speed of the reflecting wall.

Result: the radiation has the functional form (1.3).

By integration of Wien's law, one re-obtains Stefan-Boltzmann law and from eq.(1.3) Wien displacement law follows

$$\lambda_{max} T = \text{const}$$

where λ_{max} is the wavelength where $u_\lambda(\lambda, T)$ reaches its maximum. Experimentally we have

$$\lambda_{max} T = 2.9 \cdot 10^{-3}\text{ m K} .$$

Example: The temperature of the sun is 6000 K and $\lambda_{max} = 4800\text{ Angstrom} = 4800 \cdot 10^{-10}\text{ m}$ (we are in the visible spectrum). Thus

$$6000\text{ K} \cdot 4800 \cdot 10^{-10}\text{ m} = 2.9 \cdot 10^{-3}\text{ m K} .$$

Universality allows us to replace the walls with a set of harmonic oscillators. In such a setting one can compute the average emitted power

$$\bar{w}_e = \frac{2e^2}{3c^3} |\ddot{x}|^2 = \frac{2e^2\omega^2}{3mc^3} \bar{E}_\nu; \quad e^2 = \frac{q_e^2}{4\pi\epsilon_0} \text{ (MKS)}$$

where $\bar{E}_\nu$ is the average energy of the oscillator of frequency ν and the average absorbed power

$$\bar{w}_a = \frac{\pi e^2}{3m} u_\nu(\nu, T) .$$

From the energy balance we have

$$u_\nu(\nu, T) = \frac{8\pi\nu^2}{c^3} \bar{E}_\nu \tag{1.4}$$

All the problem is then to compute $\bar{E}_\nu$.

Replacing $\bar{E}_\nu$ with the value given by the classical equipartition theorem $2 \frac{1}{2}kT$ one obtains the Rayleigh-Jeans law

$$u_\nu(\nu, T) = \frac{8\pi\nu^2}{c^3} kT$$

experimentally well verified at low frequencies but in total disagreement at high frequencies (the ultraviolet catastrophe).

The number of normal modes of the electromagnetic field in a cavity is, taking into account the two polarizations of the field

$$2 \frac{4\pi\nu^2}{c^3} d\nu V = \frac{8\pi\nu^2}{c^3} d\nu V \equiv \mathcal{N}_w . \tag{1.5}$$

As the electromagnetic field is equivalent to an ensemble of harmonic oscillators, we can reach the same result applying the classical equipartition principle to these oscillators obtaining again, and much more quickly the Rayleigh-Jeans law.

Planck hypothesis: the harmonic oscillator has energy level quantized by integers $0, \epsilon_\nu, 2\epsilon_\nu, 3\epsilon_\nu, \dots$. For the partition function we have

$$Z = \sum_n e^{-\beta\epsilon_\nu} = \frac{1}{1 - e^{-\beta\epsilon_\nu}}$$

and

$$\bar{E}_\nu = -\frac{\partial \log Z}{\partial \beta} = \frac{\epsilon_\nu}{e^{\beta\epsilon_\nu} - 1} .$$

Substituting in (1.4) and imposing Wien law we have Planck's law

$$u_\nu(\nu) = \frac{8\pi\nu^2 h\nu}{c^3(e^{h\nu/kT} - 1)}$$

from which

$$\text{Stefan constant } \sigma = \frac{2\pi^5 k^4}{15c^2 h^3}; \quad \text{displacement constant } = \lambda_M T = \frac{h}{k} \frac{c}{4.965}.$$

With the experimental values of σ and of Wien displacement constant one obtains $k = 1.38 \cdot 10^{-23} \text{ J/K}$, $h = 6.67 \cdot 10^{-34} \text{ Js}$ which in addition to the new constant h provides also the value of Boltzmann constant and thus of Avogadro's number.

Why is the spacing of the energy levels of the harmonic oscillator assumed to be constant? Let us examine the partition function for large T (small β).

$$Z = \sum_n e^{-\beta E_n} \rightarrow \int_0^\infty e^{-\beta E} \rho(E) dE$$

where $\rho(E)$ is the level density. Suppose that for $E > E_M$ the distribution be $\rho(E) = c_1 E^\alpha$ with $\alpha > -1$. For small β we have

$$Z \approx \int_0^{E_M} \rho(E) dE + c_1 \beta^{-\alpha-1} \int_0^\infty x^\alpha e^{-x} dx \approx c_1 \beta^{-\alpha-1} \Gamma(\alpha + 1)$$

from which it follows for large T

$$\bar{E} \approx \frac{\alpha + 1}{\beta} = (\alpha + 1)kT.$$

If we do not want to violate the Rayleigh-Jeans law, which works fine at high temperature, we must choose $\alpha = 0$ (constant asymptotic spacing). The simplest assumption is to extend this constant spacing to all energies.

If $\alpha < -1$ we have

$$\int_0^\infty \rho(E) dE < \infty$$

i.e. a finite number N of levels. In this case for $T \rightarrow \infty$ we have

$$\bar{E} = -\frac{\partial \log Z}{\partial \beta} \rightarrow \frac{E_1 + \dots + E_N}{N}$$

i.e. the average of the energies. Again we violate the Rayleigh-Jeans law. Similarly one finds a violation of the Rayleigh-Jeans law for $\alpha = -1$.

Example of black body. The best black body is provided by the cosmic background radiation.

The decoupling of the radiation from matter occurred at about 300 000 Yr when the universe had a volume 1000^{-3} times the present volume. The present temperature is $T \approx 2.725 \text{ K}$.

A temperature of decoupling of 3000 K is consistent with $V_2/V_1 = T_1^3/T_2^3 = 1000^3$ due to an adiabatic expansion i.e. conservation of entropy, see eq.(1.2).

Keep in mind that the cosmic background radiation has to be corrected by the motion of the earth of about 600 Km/sec.

1.2 Wien's law and the number of photons

If we accept that the black body radiation is a collection of photons, the number of photons is given according to Wien's law by

$$dn = \frac{1}{h} \left(\frac{\nu}{T}\right)^2 f\left(\frac{\nu}{T}\right) VT^3 \frac{d\nu}{T}$$

i.e.

$$N = \frac{VT^3}{h} \int_0^\infty x^2 f(x) dx$$

which under an adiabatic transformation $VT^3 = \text{const}$, see eq.(1.2), is left unchanged i.e. the number of photons remains constant during the adiabatic compression or expansion.

The Doppler effect

$$\delta\nu = 2\frac{v}{c}\nu \cos \theta$$

with $v \ll c$ the speed of the moving mirror, is consistent with a momentum of the photon $p = \frac{h\nu}{c}$ and an energy of the photon $E = h\nu$. In fact the work performed by the mirror is

$$L = \int F_x v dt = v \int F_x dt = v \Delta p_x = v 2\frac{h\nu}{c} \cos \theta = h\delta\nu = \Delta E .$$

1.3 Energy fluctuations in the black body

The average value of the square of the energy fluctuations is given by

$$-\frac{\partial \bar{E}}{\partial \beta} = \overline{(E^2)} - (\bar{E})^2 = \overline{(\Delta E)^2} = kT^2 C_V . \quad (1.6)$$

Let us consider a frequency interval $\Delta\nu$. Starting from Planck's law

$$\bar{E} = \frac{8\pi\nu^2 h\nu}{c^3(e^{h\nu/kT} - 1)} \Delta\nu V \equiv h\nu \mathcal{N}_p$$

where $\mathcal{N}_p$ is the number of photons in the photon interpretation of the radiation, we obtain using (1.6)

$$\overline{(\Delta E)^2} = h\nu \bar{E} + \frac{c^3}{8\pi\nu^2 \Delta\nu V} \bar{E}^2 = (h\nu)^2 \mathcal{N}_p + \frac{\bar{E}^2}{\mathcal{N}_w}$$

where $\mathcal{N}_w$ is given in eq.(1.5).

The first term is the typical particle fluctuation with $\mathcal{N}_p$ independent systems $(\Delta E)^2 = \varepsilon_0 E \approx \varepsilon_0^2 \mathcal{N}_p$. The second part does not contain h and as such is a classical wave contribution to the energy fluctuations. The number of classical waves in the volume V , and in the frequency interval $\Delta\nu$ is given by

$$\mathcal{N}_w = \frac{8\pi\nu^2 \Delta\nu V}{c^3}$$

and if we write

$$\bar{E} = \epsilon(\nu) \mathcal{N}_w$$

we expect a fluctuation

$$\Delta E = \epsilon(\nu) \sqrt{\mathcal{N}_w}$$

i.e.

$$\overline{(\Delta E)^2} = \epsilon(\nu)^2 \mathcal{N}_w = \bar{E}^2 / \mathcal{N}_w.$$

The presence in the fluctuation of the energy of two terms, one particle-like and the other wave-like is the first evidence of the dual nature of light.

References

- [1] M. Born *Atomic physics*, Blackie and Son Ltd, Glasgow (1958)

Chapter 2

The superposition principle

2.1 Observables, spectrum, eigenstates and probability

1. Two states of the same physical system, produced by two different preparing apparatus will be considered the same state if they give rise to the same statistical distributions of results when measuring any measurable quantity.
2. We assume that it is meaningful to consider linear superposition of two or more states with complex coefficients; in different words that the space of states is a vector space on the complex numbers. This is supported by the success of the wave description of matter.
3. Given a measurable quantity we shall call spectrum of such a quantity the set of all possible results of the measurement of such a quantity on all states of the system. Such a spectrum is assumed to be composed of real numbers; complex spectra would correspond to the “simultaneous measurement” of two real numbers, an occurrence we want to exclude as we know that the measurement of whatever quantity can alter in an unpredictable way the state of the system. For simplicity sake we shall start with the discrete spectrum. We shall call eigenstate of a measurable quantity a state such that performing on it the measure of such a quantity one obtains with certainty a given value of the spectrum. If the set of

the eigenstates of a certain measurable quantity is a complete set in the space of the states we shall call such measurable quantity an observable.

4. Probabilistic interpretation of the state vectors.

Generalizing Born probabilistic interpretation of the wave function we postulate that the probability $P_A^{(i)}(\psi)$ to find the i -th value of the spectrum measuring the observable A on the state ψ is given by

$$P_A^{(i)}(\psi) = (\psi, K_A^{(i)} \psi) \quad (2.1)$$

where $K_A^{(i)}$ is a linear operator. From this it follows that ψ e $e^{i\alpha}\psi$ represent the same state.

If we consider the vector $\psi' = \rho\psi$ we have that

$$P_A^{(i)}(\psi) = (\psi', \frac{K_A^{(i)}}{|\rho|^2} \psi')$$

for any A and whatever i .

Thus we can say that the set of vectors $\rho\psi$, $\rho \neq 0$ (ray in Hilbert space) describe the same state. It is useful to work with normalized vectors.

To the null vector we cannot associate a state because we cannot compute from it any probability.

Notice that if ψ_1 is eigenstate to the value a_1 e ψ_2 is eigenstate to the value a_2 different from a_1 the two states are independent. This because if the two vectors were proportional they would give the same probabilities and thus they would be experimentally indistinguishable, while being eigenstates to different eigenvalues they are distinguishable. More generally the vector $\psi = c_1\psi_1 + c_2\psi_2$ never will be the null vector for c_1 e c_2 both different from zero and ψ_1 e ψ_2 representing different states.

2.2 The superposition principle

We shall assume the following “superposition principle”:

Given a state ψ , combination of eigenstates ψ_k of the observable A of spectrum $\{a_k\}$, the possible results of the measure of A on ψ are the a_k relative to the ψ_k which appear in the expansion of ψ .

It follows that if ψ_1 e ψ'_1 are vectors representing eigenstates of A to the same values a_1 we have that measuring A on $\psi = c_1\psi_1 + c'_1\psi'_1$ we shall find with certainty the value a_1 and thus ψ is an eigenstate of A to the value a_1 . Thus we can speak of eigensubspaces of an observable A relative to a given value of the spectrum.

Being the set of the eigenstates of A complete, every state vector can be written as

$$\psi = \sum_i c_i \psi_i \quad (2.2)$$

with ψ_i eigenstates of A all relative to different values of the spectrum.

This because we can collect all vectors which represent eigenstates relative to the same eigenvalue in a single eigenstate. Notice also that a combination $\psi = \sum_i c_i \psi_i$ with c_i not all equal to zero cannot be the null vector; otherwise we would have

$$\psi_1 = -\frac{1}{c_1} \sum_{i>1} c_i \psi_i$$

which is contradictory as, from the r.h.s. measuring A on ψ_1 we should find only values different from a_1 . Thus the states ψ_i are linearly independent.

We pose now the problem to express $P_A^{(i)}(\psi)$ in term of the coefficients of the expansion (2.2) with normalized ψ_i .

To this end we keep the ψ_i fixed and we vary properly the c_i .

According to our axiom we have

$$P_A^{(i)}(\psi) = (\psi, K_A^{(i)} \psi)$$

and thus

$$P_A^{(1)}(\psi) = \sum_{lm} c_l^* A_{lm}^{(1)} c_m.$$

Choosing c_1 equal to 1 and the others zero we derive for $l = 1$ $A_{11}^{(1)} = 1$ and for $l \neq 1$ $A_{ll}^{(1)} = 0$.

Choosing the vector obtained by normalizing $\phi = \varepsilon\psi_1 + \psi_2$ i.e.

$$\psi = \frac{\varepsilon}{N}\psi_1 + \frac{1}{N}\psi_2; \quad N^2 = (\phi, \phi)$$

and imposing

$$P_A^{(1)}(\psi) = \frac{\varepsilon^* \varepsilon}{N^2} + \frac{\varepsilon^*}{N^2} A_{12}^{(1)} + A_{21}^{(1)} \frac{\varepsilon}{N^2} \geq 0; \quad \forall \varepsilon$$

we have $A_{12}^{(1)} = A_{21}^{(1)} = 0$. Considering the vector obtained by normalizing $\phi = b_2\psi_2 + b_3\psi_3$ with b_2 e b_3 not both equal to zero and imposing

$$P_A^{(1)}(\psi) = \frac{b_2^* A_{23}^{(1)} b_3 + b_3^* A_{32}^{(1)} b_2}{N^2} = 0$$

we have $A_{23}^{(1)} = A_{32}^{(1)} = 0$.

Thus we reached the conclusion

$$P_A^{(i)}(\psi) = c_i^* c_i .$$

As a corollary we have

$$\sum_i c_i^* c_i = \sum_i P_A^{(i)}(\psi) = 1$$

as with certainty we obtain some result. From this the orthogonality property follows. Let us consider in fact $\phi = \sum_j b_j \psi_j$, with b_j not all vanishing. We know that such vector is not the null vector. If we denote by N its norm the normalized vector is $\psi = \sum_j c_j \psi_j$ with $c_j = \frac{1}{N} b_j$. Thus we have

$$N^2 = \sum_{lm} b_l^* (\psi_l, \psi_m) b_m = N^2 \sum_k c_k^* c_k = \sum_k b_k^* b_k .$$

This identity in the b_k tells us that $(\psi_l, \psi_m) = \delta_{lm}$. Thus we reach the fundamental result: Eigenvectors of the same observable relative to different values of the spectrum are orthogonal.

2.3 Standard expansion of a vector

In each eigensubspace relative to the value a_k of the spectrum of A one chooses a complete orthonormal basis $\psi_k^{(i)}$. Every vector can be written in the form

$$\psi = \sum_{ki} c_k^{(i)} \psi_k^{(i)} = \sum_k c_k \psi_k$$

with

$$\sum_i c_k^{(i)} \psi_k^{(i)} = c_k \psi_k \quad \text{from which} \quad |c_k|^2 = \sum_i |c_k^{(i)}|^2 .$$

It follows that

$$P_A^{(k)}(\psi) = \sum_i |c_k^{(i)}|^2 .$$

2.4 Observable with bounded but otherwise arbitrary spectrum

We refer in this section to an observable with bounded but otherwise arbitrary spectrum (point, continuous, mixed ...). The treatment will be rigorous.

We divide the spectrum in a finite number of intervals $[a_{i-1}, a_i]$, $i = 1, \dots, n$.

The Born rule (2.1) is generalized to

$$\text{Prob}_A[a_{i-1}, a_i] = (\psi, K_A[a_{i-1}, a_i]\psi)$$

and the superposition principle to: If measuring A on the state ϕ_j we obtain with certainty values in the interval $[a_{k_j-1}, a_{k_j}]$, the only values we can obtain by measuring A on the state

$$\sum_j c_j \phi_j$$

are contained in the union of the intervals $[a_{k_j-1}, a_{k_j}]$.

If A is an observable every vector can be obtained as the sum of vectors $\psi[a_{i-1}, a_i]$ i.e.

$$\psi = \sum_{i=1}^n \psi[a_{i-1}, a_i]$$

with the property that measuring A on $\psi[a_{i-1}, a_i]$ (when $\psi[a_{i-1}, a_i] \neq 0$) we are sure to find a value in $[a_{i-1}, a_i]$.

The consequence is that the non zero vectors $\psi[a_{i-1}, a_i]$ are independent. Otherwise for some k we could write $\psi[a_{k-1}, a_k]$ as a combination of others. Then from the l.h.s. we would be sure to measure a value in the interval $[a_{k-1}, a_k]$ while from the r.h.s. we would be sure to find a value outside $[a_{k-1}, a_k]$. Thus we have also that the decomposition

$$\psi = \sum_{i=1}^n \psi[a_{i-1}, a_i]$$

is unique. The operation $\psi \rightarrow \psi[a_{i-1}, a_i]$ is linear and the decomposition of $\psi[a_{i-1}, a_i]$ is just itself, i.e. the operation $\psi \rightarrow \psi[a_{i-1}, a_i] \equiv P_i \psi$ is a projector $P_i = P_i^2$. Repeating the reasoning used in the discrete case we obtain

a) $\text{Prob}_A[a_{i-1}, a_i] = (\psi[a_{i-1}, a_i], \psi[a_{i-1}, a_i])$

b) $(\psi[a_{j-1}, a_j], \psi[a_{i-1}, a_i]) = 0$ for $i \neq j$

i.e. P_i are orthogonal projectors. In fact

$$0 = (P_i \psi, (1 - P_i) \psi) = (\psi, P_i^+ (1 - P_i) \psi)$$

$P_i^+ = P_i^+ P_i$ i.e. $P_i = P_i^+$, $P_i P_j = 0$ for $i \neq j$.

Thus the average value of the measures of A is

$$\bar{a} = \lim_{\max \Delta_i \rightarrow 0} \sum_i a'_i (\psi[a_{i-1}, a_i], \psi[a_{i-1}, a_i]) = \lim_{\max \Delta_i \rightarrow 0} \sum_i a'_i (\psi, P_i \psi) \quad \text{with } a_{i-1} \leq a'_i < a_i$$

where $\Delta_i = a_i - a_{i-1}$.

The existence of such a limit is obtained by comparing two elements with a different choice of intervals with one choice which belongs to the refinement of the two sets of intervals and using

$$\sum_i (\psi[a_{i-1}, a_i], \psi[a_{i-1}, a_i]) = 1 .$$

We construct the operator $\hat{A}$ by

$$\hat{A}\psi \equiv \lim_{\max \Delta_i \rightarrow 0} \sum_i a'_i \psi[a_{i-1}, a_i] .$$

Such limit exists. The proof is obtained by comparing two element with one which belongs to the refinement of the two sets of intervals and noticing that due to orthogonality we have

$$\| \sum_i c_i \psi[a_{i-1}, a_i] \| \leq \max(|c_i|) \|\psi\| . \quad (2.3)$$

We can define $P_i = E(a_i) - E(a_{i-1})$ with $E(a) = P[-\infty, a)$ and thus

$$\hat{A} = \lim_{\max \Delta_i \rightarrow 0} \sum_i a'_i (E(a_i) - E(a_{i-1})) \equiv \int a dE(a) \quad (2.4)$$

where due to (2.3) the convergence is uniform. $\hat{A}$ is a bounded hermitean operator as uniform limit of hermitean operators and we have

$$\bar{a} = (\psi, \hat{A}\psi).$$

The probability of finding a result in the interval $[a_{j-1}, a_j]$ is given by

$$\text{Prob}_A[a_{j-1}, a_j] = (\psi, (E(a_j) - E(a_{j-1}))\psi) .$$

Every self-adjoint operator, also unbounded, can be written in the form (2.4) [1].

Thus we see that due to the Born rule the value of the average of the measurements of the observable A is given by the mathematical mean value of the operator $\hat{A}$ associated to the observable A . This is also trivially true for the average value of any function of the observable A with the usual mathematical definition for the function of a self-adjoint operator A

$$A = \int a \, dE(a), \quad f(A) \equiv \int f(a) dE(a) .$$

A different path to set up the mathematical scheme is, following Dirac, to assume a priori that to each observable there correspond a self-adjoint operator and that the mathematical mean value of the operator $f(A)$

$$(\psi, f(A)\psi)$$

gives the experimental average value of the function $f(a)$ for any f . From this Born rule follows. It is in fact sufficient to work with the function $f(a)$ which equals 1 for $b_1 \leq a < b_2$ and zero everywhere else, for any choice of b_1 and b_2 . Then such a mean value is the probability to measure a value of a in the interval $[b_1, b_2)$.

It has to be noticed that the characterization of an observable is its spectral family more than its spectrum. In fact as far as the spectrum is concerned if we consider instead of A , a function of A e.g. $\tanh A$, the spectrum of such operator, discrete or not, is contained in the interval $[-1, 1]$. The bounded operator $\tanh A$ is perfectly defined from the spectral decomposition of A . This does not mean that we can use always observables represented by bounded operators. E.g. Stone theorem relates the group of evolution operators $U(t)$ with the hamiltonian H which is its generator and in most cases such generator is an unbounded, self-adjoint, operator (see Chapters IV and V). Working with $\tanh H$ would be completely unpractical.

2.5 Continuum spectrum in the Dirac formalism

In this paragraph we give without any pretense of rigor the treatment of an observable with a non degenerate continuous spectrum in the formalism of Dirac i.e. employing a vector space larger than a Hilbert space.

The rigorous treatment of such a space (the Gelfand triplet) is a very elaborate subject [2]. On the whole, for a rigorous treatment of a problem, the spectral family treatment followed in the previous section is simpler while Dirac's formalism is more elegant and more inspiring.

Given a physically measurable quantity Q with continuous spectrum such quantity will be called observable if

1) For every point q of the spectrum a vector $|q\rangle$ is associated such that every vector of the Hilbert space (or better a proper dense subset of vectors of the Hilbert space) can be written as a continuous superposition of vectors $|q\rangle$

$$|\varphi\rangle = \int c(q)|q\rangle dq . \quad (2.5)$$

2) If a vector $|\varphi_I\rangle$ can be written as an integral on the interval I

$$|\varphi_I\rangle = \int_I c(q)|q\rangle dq$$

the only values of the spectrum of Q which can be obtained measuring Q on $|\varphi_I\rangle$ are contained in the interval I .

Considering now finite linear combinations of vectors $|\varphi_{I_1}\rangle$, $|\varphi_{I_2}\rangle$ etc. relative to intervals with zero intersection and reasoning exactly as in the case of the discrete spectrum we obtain the following results

a) given the normalized vector $|\varphi\rangle$ i.e. $\langle\varphi|\varphi\rangle = 1$

$$|\varphi\rangle = \int c(q)|q\rangle dq$$

the probability to find, measuring Q on $|\varphi\rangle$, values contained in the interval I is given by $\langle\varphi_I|\varphi_I\rangle$ with

$$|\varphi_I\rangle = \int_I c(q)|q\rangle dq$$

b) Vectors $|\varphi_{I_j}\rangle$ relative to intervals with zero intersection are orthogonal.

As for $I_1 \cap I_2 = \emptyset$ we have

$$\langle \varphi_{I_2} | \varphi_{I_1} \rangle = \int_{I_2} \int_{I_1} c_2^*(q) \langle q' | q \rangle c_1(q) dq' dq = 0$$

for arbitrary $c_1(q)$ e $c_2(q)$, we have that $\langle q' | q \rangle = 0$ for $q' \neq q$.

If we assume that $\langle q' | q \rangle$ is a function we reach the paradox that the norm of every vector (2.5) is zero, as the integrand appearing in the scalar product is different from zero only on the zero measure set $q' = q$. This shows the necessity to widen the Hilbert space, if we insist in dealing with eigenstates of an observable with continuum spectrum.

The way out is to admit that $\langle q' | q \rangle$ be a distribution, in particular

$$\langle q' | q \rangle = k(q) \delta(q' - q) .$$

The function $k(q)$ must be positive otherwise we violate the positivity of the norm in Hilbert space. Thus we can normalize the vectors $|q\rangle$ dividing by $\sqrt{k(q)}$ obtaining the standard normalization.

$$\langle q' | q \rangle = \delta(q' - q) . \quad (2.6)$$

Using the standard normalization we have for the probability to find a value of the spectrum in the interval I

$$P_I = \langle \varphi_I | \varphi_I \rangle = \int_I c^*(q) c(q) dq$$

from which it follows that the probability density is given by

$$P(q) = c^*(q) c(q) .$$

Exploiting (2.6) we have

$$c(q) = \langle q | \varphi \rangle$$

and thus

$$|\varphi\rangle = \int |q\rangle \langle q | \varphi \rangle dq$$

from which the completeness relation

$$\int |q\rangle \langle q| dq = \mathbf{1}$$

follows.

In completely analogous manner as done in the case of the discrete spectrum we can associate to the observable Q the operator $\hat{Q}$ given by

$$\hat{Q} = \int q |q\rangle \langle q| dq .$$

We have

$$\langle \varphi | \hat{Q} | \varphi \rangle = \int q P(q) dq = \bar{q} .$$

For any polynomial p we have

$$p(\hat{Q}) = \int p(q) |q\rangle \langle q| dq$$

which is extended to any function f defined on the spectrum of Q by

$$f(\hat{Q}) = \int f(q) |q\rangle \langle q| dq$$

References

- [1] F. Riesz and B. Nagy *Functional Analysis*, Dover Publications, Inc. New York (1990)
- [2] I.M. Gelfand and N. Ya. Vilenkin *Generalized functions*, Vol 5, Academic press, New York, London (1964)

Chapter 3

The correspondence principle

We look in the space of operators for an operation which has the same algebraic properties of the classical Poisson brackets, i.e. a binary operation (A, B) which induces on the space of operators a Lie algebra [1,2] i.e. an operation which is linear both in A and B and antisymmetric in A and B and such that

$$(AB, C) = A(B, C) + (A, C)B$$

i.e. the Leibniz rule.

Expanding

$$(A_i A_j, B_i B_j)$$

in two different orders and using Leibniz rule one obtains [1]

$$[A_i, B_i](A_j, B_j) = (A_i, B_i)[A_j, B_j]$$

where the square brackets denote the commutator, for whatever choice of the operators A_k, B_k . Defined

$$[L] = \sum_i a_i [A_i, B_i]$$

$$(L) = \sum_i a_i (A_i, B_i)$$

one obtains

$$[L](A_j, B_j) = (L)[A_j, B_j]$$

$$L = (L)[L] .$$

We admit that there exists an $[L]$ which possesses the inverse $[L]^{-1}$.

It follows that

$$(A_j, B_j) = [A_j, B_j](L)[L]^{-1}$$

$$(A_j, B_j) = [L]^{-1}(L)[A_j, B_j] .$$

But $[L]^{-1}(L) = (L)[L]^{-1}$ and thus $(L)[L]^{-1}$ commutes with all commutators and thus with the algebra generated by all commutators. A set $\mathcal{M}$ of operators is said irreducible if there exists no proper subspace of $\mathcal{H}$ invariant under the action of $\mathcal{M}$.

Let $\mathcal{M}$ be an algebra of bounded operators on the complex field such that if $A \in \mathcal{M}$ also $A^+ \in \mathcal{M}$.

The following three statements are equivalent [3]

1. $\mathcal{M}$ is irreducible.
2. The commutant of $\mathcal{M}$ (i.e. $\mathcal{M}'$) is cI .
3. Every vector $\psi \neq 0$ is cyclic i.e. $\overline{\text{linear span}(\mathcal{M}\psi)} = \mathcal{H}$.

This is a generalization of the classical Schur Lemma.

Thus if the algebra $\mathcal{M}$ generated by commutators is irreducible, i.e. sufficiently large, we have $\mathcal{M}' = \{\lambda I\}$, and then

$$(A_j, B_j) = c[A_j, B_j]$$

and c is a universal constant. If we want to respect the same dimensional relations as in the classical Poisson brackets, c must have the dimension of an action. Moreover imposing that (A, B) with A e B hermitean be hermitean, we have that c has to be pure imaginary.

We set

$$(A_j, B_j) = \frac{1}{i\hbar}[A_j, B_j] \tag{3.1}$$

where $\hbar$ is a constant with the dimensions of an action.

We warn however that while classical mechanics is invariant under canonical transformations and in particular under canonical transformations induced by lagrangian transformations, the translation (3.1) of the Poisson brackets into commutators is not unique due to the ordering of non commuting operators.

The choice of cartesian coordinates turns out to be the correct one i.e. the one consistent with experimental data. The constant $\hbar$ has also to be determined experimentally.

We note that statement (3) tells us that in case of irreducibility we can replace the vectors with operators; to superpose two states $A\psi$ and $B\psi$ is equivalent to consider $\alpha A + \beta B$ and thus the vector space of states is actually the linear space of operators.

Example 1.

Let us consider the Hilbert space of the L^2 functions $f(\mathbf{q})$ (space of the wave functions describing a single spinless particle). Consider on such a space the operators $\hat{q}_k$ given by the multiplication by q_k and the operators $\hat{t}_j$ given by the action of $-i\frac{\partial}{\partial q_j}$. We have

$$[\hat{q}_k, \hat{t}_j] = i\delta_{kj} .$$

Thus we have one commutator $[L]$ (actually more than one) which admits inverse $[L]^{-1}$ and $(L)[L]^{-1}$ will commute with the whole algebra generated by commutators. We have

$$[\hat{q}_k^n, \hat{t}_j] = in\hat{q}_k^{n-1}\delta_{kj}$$

$$[\hat{q}_k, \hat{t}_j^m] = im\hat{t}_j^{m-1}\delta_{kj}$$

i.e. the algebra of commutators contains all powers on $\hat{q}_k$ and all powers of $\hat{t}_j$.

Consider now an operator F which commutes with all $\hat{q}_k$. We have with

$$\langle \mathbf{q}' | F | \mathbf{q} \rangle = f(\mathbf{q}', \mathbf{q})$$

$$0 = \langle \mathbf{q}' | [\hat{q}_k, F] | \mathbf{q} \rangle = (q'_k - q_k) f(\mathbf{q}', \mathbf{q})$$

i.e.

$$\langle \mathbf{q}' | F | \mathbf{q} \rangle = \delta(\mathbf{q}' - \mathbf{q}) f(\mathbf{q}) = \delta(\mathbf{q}' - \mathbf{q}) f(\mathbf{q}') .$$

If now F commutes with $\hat{t}_j$ we have

$$0 = \langle \mathbf{q}' | [\hat{t}_j, F] | \mathbf{q} \rangle = -i\delta(\mathbf{q}' - \mathbf{q}) \frac{\partial f(\mathbf{q}')}{\partial q'_j}$$

from which we have $f(\mathbf{q}) = \text{const.}$ and thus we proved statement (2). Thus the only binary operation on operators on the considered Hilbert space which

possesses the algebraic properties of Poisson brackets is the commutator. In particular the only translation of the Poisson brackets to operators on the Hilbert space of the L^2 functions of three variables is

$$[\hat{q}_j, \hat{q}_k] = 0, \quad [\hat{p}_j, \hat{p}_k] = 0, \quad [\hat{q}_j, \hat{p}_k] = i\hbar \delta_{jk}, \quad (j, k = 1, 2, 3) \quad (3.2)$$

where $\hbar$ is a constant to be determined experimentally e.g. by diffraction of electrons on crystals (there are also more precise ways of identifying $\hbar$). Following the formalism employed above one can easily prove von Neumann theorem: All realization of the commutation relations (3.2) are unitary equivalent. In fact after writing

$$\hat{p}_k = -i\hbar \frac{\partial}{\partial q_k} + \hat{R}_k$$

the last of eq.(3.2) tells us

$$\langle \mathbf{q}' | \hat{R}_k | \mathbf{q} \rangle = \delta(\mathbf{q}' - \mathbf{q}) r_k(\mathbf{q})$$

i.e.

$$\hat{p}_k = -i\hbar \frac{\partial}{\partial q_k} + r_k(\mathbf{q}) .$$

The second of eq.(3.2) tells us that

$$\frac{\partial r_k(\mathbf{q})}{\partial q_j} = \frac{\partial r_j(\mathbf{q})}{\partial q_k}$$

i.e. for a simply connected space

$$r_k(\mathbf{q}) = \frac{\partial}{\partial q_k} \Lambda(\mathbf{q})$$

$\Lambda(\mathbf{q})$ real. Perform now the unitary transformation on the basis $|\mathbf{q}\rangle$

$$|\mathbf{q}\rangle = e^{i\Lambda/\hbar} |\mathbf{q}\rangle .$$

We have

$$\begin{aligned} \langle \mathbf{q} | \hat{p}_k | \psi \rangle &= e^{i\frac{\Lambda}{\hbar}} \left(-i\hbar \frac{\partial}{\partial q_k} + \frac{\partial \Lambda}{\partial q_k} \right) \langle \mathbf{q} | \psi \rangle = e^{i\frac{\Lambda}{\hbar}} \left(-i\hbar \frac{\partial}{\partial q_k} + \frac{\partial \Lambda}{\partial q_k} \right) e^{-i\frac{\Lambda}{\hbar}} \langle \mathbf{q} | \psi \rangle \\ &= -i\hbar \frac{\partial}{\partial q_k} \langle \mathbf{q} | \psi \rangle . \end{aligned}$$

For a rigorous proof of von Neumann theorem starting from Weyl algebra

$$e^{ipa/\hbar} e^{iqb} e^{-ipa/\hbar} = e^{iqb} e^{iab}$$

which has the advantage of dealing only with bounded operators, see [4].

Example 2.

Let us consider the two-dimensional Hilbert space which describes the spin i.e. angular momentum 1/2, with the operators

$$s_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad s_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix};$$

$$[s_j, s_k] = i\varepsilon_{jkm} s_m.$$

The operators s_1, s_2, s_3 plus the identity form a complete set of matrices and thus we have that the algebra generated by the commutators is irreducible. It follows that the unique operation which respects the properties of the Poisson brackets is the commutator.

References

- [1] P.A.M. Dirac *The principles of quantum mechanics*, Chap.4, Clarendon Press, Oxford (1958)
- [2] J. Grabowski e G. Marmo *Binary operations in classical and quantum mechanics*, Banach Center Publications, vol. 59, Polish Academy of Sciences (2003)
- [3] W. Thirring *Quantum mechanics of atoms and molecules*, Chap.2, Springer Verlag (1981)
- [4] W. Thirring *Quantum mechanics of atoms and molecules*, Chap.3, Springer Verlag (1981)

Chapter 4

The time evolution

4.1 Introduction

For the time evolution operator $U(t_2, t_1)$ we shall require:

- 1) Linearity, which we consider a fundamental feature of quantum mechanics.
- 2) Range $U = \mathcal{H}$, i.e. we can produce at time t_2 whatever state vector provided we start at time t_1 with a proper vector.
- 3) Isometry of $U(t_2, t_1)$. This is not an optional requirement if we want to maintain linearity. In fact if we try to normalize the transformation

$$U\phi_1 = \phi'_1, \quad U\phi_2 = a\phi'_2, \quad U(\phi_1 + c\phi_2) = \phi'_1 + ca\phi'_2$$

with $|a| \neq 1$ and the ϕ 's, ϕ' 's normalized, by going over to U'

$$U'\phi_1 = e^{i\alpha_1}\phi'_1, \quad U'\phi_2 = e^{i\alpha_2}\phi'_2, \quad U'(\phi_1 + c\phi_2) = b(\phi'_1 + ca\phi'_2)$$

we would have from the linearity of U'

$$U'(\phi_1 + c\phi_2) = e^{i\alpha_1}\phi'_1 + ce^{i\alpha_2}\phi'_2 = b(\phi'_1 + ca\phi'_2)$$

giving

$$e^{i\alpha_1} = ae^{i\alpha_2}$$

which cannot hold for $|a| \neq 1$.

From 2) and 3) it follows that $U(t_2, t_1)$ is unitary. In fact the isometry of U gives us $U^+U = 1$ and due to $\text{Range } U = \mathcal{H}$ given $\zeta \in \mathcal{H}$, a η exists such that

$$\zeta = U\eta .$$

Then

$$U^+\zeta = \eta, \quad UU^+\zeta = \zeta .$$

Thus U has an inverse which equals U^+

$$U^+(t_2, t_1)U(t_2, t_1) = U(t_2, t_1)U^+(t_2, t_1) = I .$$

The fact that it is physically meaningful to superpose two states with complex coefficient and that such superposition, due to the unitary nature of the time evolution, is maintained as the states evolve is the key feature of quantum mechanics.

4.2 Composition property of U

We shall denote the inverse of $U(t_2, t_1)$ by $U(t_1, t_2) \equiv U(t_2, t_1)^{-1}$. In the case of invariance under time translation we have

$$U(t_1 + \tau, t_1) = U(t_2 + \tau, t_2)$$

from which

$$U(t_2, t_1) = U(t_2 - t_1, 0) \equiv V(t_2 - t_1) = V(-t_2 + t_1)^{-1}$$

$$V(t_1)V(t_2) = V(t_2)V(t_1) = V(t_1 + t_2) .$$

Thus we have a one parameter group of unitary transformations. We shall require weak continuity in t .

We have now the fundamental

Stone theorem: If a one parameter abelian group of unitary transformations is weakly continuous it can be written as the exponential of a self-adjoint operator.

$$U(t) = \exp(-iAt)$$

and we have

$$\left. \frac{dU(t)}{dt} \right|_{t=0} = -iA \quad (4.1)$$

in the sense that for all and only all $\psi \in D(A)$ we have

$$\lim_{\epsilon \rightarrow 0} \frac{U(\epsilon) - U(0)}{\epsilon} \psi = -iA\psi.$$

Vice-versa given a self-adjoint operator A we have that $U(t) = \exp(-iAt)$ is a one parameter unitary group weakly continuous and eq.(4.1) holds.

We have

$$\frac{dU(t)}{dt} = -iAU(t) = -iU(t)A.$$

In fact

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{U(t+\epsilon) - U(t)}{\epsilon} \psi &= \lim_{\epsilon \rightarrow 0} \frac{U(\epsilon) - U(0)}{\epsilon} U(t) \psi = \\ &= U(t) \lim_{\epsilon \rightarrow 0} \frac{U(\epsilon) - U(0)}{\epsilon} \psi = -iU(t)A\psi \end{aligned}$$

which shows that if $\psi \in D(A)$ we have also $U(t)\psi \in D(A)$.

There exists an other important correspondence between self-adjoint operators and unitary operators given by the Cayley transformation

$$U = (A - iI)(A + iI)^{-1}$$

and its inverse

$$A = i(I + U)(I - U)^{-1}$$

where U is a unitary transformation with 1 not a characteristic value.

Coming back to (4.1) the evolution equation becomes

$$i\hbar \frac{d\psi}{dt} = H\psi$$

with H self-adjoint operator with the dimensions of energy. The picture in which the evolution of the physical system is described by the change in time of the state vector is called the Schrödinger picture.

In absence of invariance under time translation we have, in the simple case where $U(t, t_1)$ admits a bounded derivative

$$\frac{dU(t, t_1)}{dt} = -iA(t, t_1)U(t, t_1) .$$

But as

$$\frac{dU(t, t_1)}{dt} = -iA(t, t_2)U(t, t_2)U(t_2, t_1) = -iA(t, t_2)U(t, t_1)$$

we have that A depends only on t .

$$\frac{dU(t, t_1)}{dt} = -iA(t)U(t, t_1).$$

It is easily seen that the unitarity of U imposes $A(t) = A^+(t)$. If $A(t) = A$ after constructing $U(t) = \exp(-iA(t - t_1))$ one shows that $U^+(t)U(t, t_1) = \text{const}$ i.e. $U(t, t_1) = \exp(-iA(t - t_1))$. The general time-dependent case in which $A(t)$ is not bounded is rather difficult to handle as even the domain of $A(t)$ itself may change in time [2].

The reason why $A(t)$ is so important is that the description of physics local in time is simple while in general the solutions for finite time are complicated.

We must now physically identify the operator $A(t)$. The Heisenberg picture is the most suitable one

$$(\phi(t), B(t)\psi(t)) = (\phi(t_0), B_H(t)\psi(t_0))$$

with

$$B_H(t) = U^+(t, t_0)B(t)U(t, t_0)$$

and

$$\frac{dB_H(t)}{dt} = \frac{\partial B_H(t)}{\partial t} + \frac{1}{i\hbar}[B_H(t), H_H(t)]$$

which also defines $\frac{\partial B_H(t)}{\partial t}$. As the commutation relations are invariant under unitary transformations we have also

$$[q_{H_j}(t), q_{H_l}(t)] = 0, \quad [p_{H_j}(t), p_{H_l}(t)] = 0, \quad \frac{1}{i\hbar}[q_{H_j}(t), p_{H_l}(t)] = \delta_{jl}$$

and comparing with classical mechanics where

$$\frac{dB(t)}{dt} = \frac{\partial B(t)}{\partial t} + [B(t), H(t)]_{PB}$$

we have that $H_H(t)$ is naturally identified with the Hamiltonian.

4.3 Construction of the solution to the time evolution equation

In the case in which H is independent of time the problem is solved by Stone theorem.

If $H(t)$ is constant on intervals we have

$$U(t_n, t_0) = \Pi_j e^{-\frac{i}{\hbar} H_j(t_j - t_{j-1})}$$

If $H(t)$ is properly approximated by an H constant over intervals we have [2]

$$U(t_n, t_0) = \lim_{\max \Delta_j \rightarrow 0} \Pi_j e^{-\frac{i}{\hbar} H_j(t_j - t_{j-1})} \equiv \Pi e^{-\frac{i}{\hbar} H(t) dt}.$$

4.4 Expansion in T -ordered integrals

We start again from

$$i\hbar \frac{\partial U(t_2, t_1)}{\partial t_2} = H(t_2) U(t_2, t_1). \quad (4.2)$$

To give an idea of the expansion in T -ordered integrals we refer to the case in which $H(t)$ is bounded and uniformly continuous and its derivative exists in uniform sense. This occurs in a number of interesting cases. The most important one occurs in perturbation theory. In that case the Hamiltonian is replaced by $\tilde{V}(t) = e^{i\frac{H_0 t}{\hbar}} V(t) e^{-i\frac{H_0 t}{\hbar}}$. If $V(t)$ is hermitean, as it must be, and bounded, also $\tilde{V}(t)$ is hermitean and bounded.

After setting $U(t_2, t_1) = I + K(t_2, t_1)$ we write the Volterra equation

$$K(t_2, t_1) = -\frac{i}{\hbar} \int_{t_1}^{t_2} H(t') dt' - \frac{i}{\hbar} \int_{t_1}^{t_2} H(t') K(t', t_1) dt'.$$

Iterating we have Dyson's series

$$U(t_2, t_1) = 1 - \frac{i}{\hbar} \int_{t_1}^{t_2} H(t') dt' + \left(-\frac{i}{\hbar}\right)^2 \int_{t_1}^{t_2} dt' \int_{t_1}^{t'} dt'' H(t') H(t'') + \dots$$

Such a series converges uniformly in norm, as also the series of its derivatives. It follows that such a series solves the starting equation.

Uniqueness:

For a solution of eq.(4.2) we have

$$\frac{d(U^+U)}{dt_2} = 0$$

i.e. $U^+U = 1$. It follows that U is isometric and thus has norm 1. Thus $K(t_2, t_1)$ has norm less or equal to 2. If there are two solutions the difference W must obey

$$W(t_2, t_1) = -\frac{i}{\hbar} \int_{t_1}^{t_2} H(t') W(t', t_1) dt'$$

which implies

$$\|W(t_2, t_1)\| \leq c \int_{t_1}^{t_2} \|W(t, t_1)\| dt \leq 4c(t_2 - t_1) .$$

Iterating we have

$$\|W(t_2, t_1)\| \leq 4(c(t_2 - t_1))^n / n!$$

i.e. $W = 0$.

Unitarity:

Reshuffling Dyson's series we have that it holds

$$-i\hbar \frac{\partial U(t_2, t_1)}{\partial t_1} = U(t_2, t_1) H(t_1) . \quad (4.3)$$

In fact e.g.

$$K_2(t_2, t_1) \equiv \left(-\frac{i}{\hbar}\right)^2 \int_{t_1}^{t_2} dt' \int_{t_1}^{t'} H(t') H(t'') dt'' = \left(-\frac{i}{\hbar}\right)^2 \int_{t_1}^{t_2} dt'' \int_{t''}^{t_2} H(t') H(t'') dt'$$

and

$$\frac{dK_2(t_2, t_1)}{dt_1} = -\left(-\frac{i}{\hbar}\right)^2 \int_{t_1}^{t_2} H(t') dt' H(t_1) .$$

From (4.3) it follows

$$\frac{d(U(t_2, t_1)U^+(t_2, t_1))}{dt_1} = 0 \quad (4.4)$$

and as for $t_1 = t_2$, $U(t_2, t_2) = 1$ we have $UU^+ = 1$.

References

- [1] H. Thirring *Quantum mechanics of atoms and molecules*, Chap.3.3, Springer Verlag (1981)
- [2] J.D. Dollard and C.N. Friedman *Product integration with applications to differential equations*, Chap.3, Addison Wesley, London (1979)

Chapter 5

Properties of the hamiltonian operator

5.1 Introduction

Often the hamiltonian operator is given in the form of a differential operator. The formal expression of a differential operator in general is not sufficient to define the operator on a Hilbert space insofar it is necessary to give its definition domain. This because due to the Hellinger-Toeplitz theorem, an operator defined on the whole Hilbert space and hermitean is necessarily bounded. On the other hand differential operators are unbounded and thus if self-adjoint they cannot be defined on the whole Hilbert space. In general the problem of the definition of a self-adjoint operator (or of the self-adjoint extension of an operator) is rather subtle. E.g. the sum of two self-adjoint operators is not necessarily self-adjoint. Let us consider for concreteness a very useful case i.e. the operator acting on the space of functions $f(x_1, x_2, x_3)$

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(r)$$

with $V(r)$ which goes to zero for $r \rightarrow \infty$ and

$$|V(r)| < \text{const. } r^{-3/2+\varepsilon} \tag{5.1}$$

for $r \rightarrow 0$. Let us assume as initial domain of definition of H the set of infinitely differentiable functions with compact support $D_0(H) = C_0^\infty$; these are obviously L^2 functions and due to (5.1) for $\phi \in D_0(H)$ it holds

$$H\phi \in \mathcal{H} .$$

We recall that once given the definition domain of H , the domain of its adjoint is completely defined and in our case it is easy to see that

$$D_0(H^+) \supset D_0(H) = C_0^\infty .$$

We need now to extend the domain $D_0(H)$ to a domain $D(H)$ such that

$$D(H^+) = D(H)$$

and such that on $D(H)$ we have $H = H^+$. Only in this case we can speak of H as a self-adjoint operator. If such a process can be carried out in more than one way we shall have more self-adjoint extensions and to each of them there corresponds a different physical problem. For self-adjoint operators one can prove the existence of the spectral representation we discussed in section (2.4). Such representation is fundamental in providing the statistical interpretation of the state vector.

To start we prove the following result: If a self-adjoint extension is possible, $D(H)$ (which coincides with $D(H^+)$) cannot contain functions which diverge at a point (e.g. the origin) like r^{-1} or faster than r^{-1} .

In fact let $\psi \in D(H)$ and $\phi \in D_0(H) = C_0^\infty \subset D(H)$. $H = H^+$ implies

$$(\psi, H\phi) = (H\psi, \phi)$$

i.e.

$$0 = \int (\psi^*(x)\nabla^2\phi(x) - \nabla^2\psi^*(x)\phi(x))d^3x \equiv \int \nabla(\psi^*(x)\nabla\phi(x) - \nabla\psi^*(x)\phi(x))d^3x .$$

Computing the last term using Gauss theorem and the fact that ϕ has compact support we have

$$0 = \lim_{r \rightarrow 0} \int_{\Sigma_r} \mathbf{n} \cdot (\psi^*(x)\nabla\phi(x) - \nabla\psi^*(x)\phi(x))d\Sigma_r .$$

But if $\psi \approx r^{-\alpha}$ the first term is $O(r^{-\alpha+2})$ while the second is $O(r^{-\alpha-1+2})\phi(0)$. Choosing $\phi(0) \neq 0$ we see that for $\alpha \geq 1$ the result is non zero or even divergent. Such a result is important to exclude e.g. in the treatment of the hydrogen atom square integrable solutions ψ which behave at the origin like $1/r$. Notice that this is due to the fact that we chose as initial domain $D_0(H) = C_0^\infty$. Had we chosen for $D_0(H)$ the functions C_0^∞ whose support excludes the origin we would have had the possibility of more than one self-adjoint extensions. Even such self-adjoint extensions, which do not give rise to Bohr spectrum for the s -wave have a physical interpretation.

5.2 Some properties of the spectrum

Let us use the trial functions

$$\psi(r) = \frac{1}{r_0^{3/2}} \frac{r_0}{r} f\left(\frac{r}{r_0}\right)$$

with $f(x) \in C_0^\infty$ with a support which exclude the origin and

$$4\pi \int_0^\infty f^2(x) dx = 1 .$$

We have

$$4\pi \int_0^\infty |r\psi(r)|^2 dr = 1, \quad T\psi = -\frac{\hbar^2}{2m} \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\psi)$$

and

$$(\psi, T\psi) = c_T \frac{1}{r_0^2}, \quad c_T > 0$$

while for $V(r) \approx \frac{c_V}{r^s}$ and small r_0 we have

$$(\psi, V\psi) \approx c_P \frac{c_V}{r_0^s}, \quad c_P > 0$$

which proves that for $s > 2$ and $c_V < 0$ the spectrum of the Hamiltonian cannot be bounded from below.

If at infinity $V \approx -c \frac{1}{r^s}$ with $s < 2$ and $c > 0$, H has an infinite number of bound states.

In fact starting from an $f(x)$ with support $(1, 2)$ we can construct an infinite sequence of state function $\psi_1, \psi_2, \psi_3 \dots$ whose wave functions have disjoint supports and on which the mean value of H is negative. This is obtained by choosing $r_0 = c$ giving the support $(c, 2c)$, then $r_o = 2c$ giving the support $(2c, 4c)$ and so on.

Thus we have $(\psi_n, \psi_m) = \delta_{mn}$ and due to the locality of the H , $(\psi_m, H\psi_n) = k_n \delta_{mn}$ with $k_n < 0$ for c sufficiently large.

Due to the completeness of the eigenvectors of H we can write

$$\psi_n = \phi^{(n)} + \zeta^{(n)}$$

where $\phi^{(n)}$ are combinations of ϕ_n and $\zeta^{(n)}$ superpositions of ζ_E with

$$H\phi_n = E_n\phi_n, \quad E_n \leq 0$$

$$H\zeta_E = E\zeta_E, \quad E > 0.$$

If the ϕ_n are finite in number N , then $\sum_{n=1}^{N+1} \alpha_n \phi^{(n)} = 0$ for some choice of α_n not all equal to zero and

$$\psi = \sum_{n=1}^{N+1} \alpha_n \psi_n = \int \beta(E) \zeta_E dE \neq 0$$

because the ψ_n are all orthogonal.

The last structure gives us $(\psi, H\psi) \geq 0$ while the first gives us $(\psi, H\psi) = \sum_n |\alpha_n|^2 (\psi_n, H\psi_n) < 0$, which is contradictory.

5.3 Lower boundedness of the spectrum

We come now to the lower boundedness of the spectrum in the case in which $|V(r)| < \text{const } r^{-\alpha}$ with $\alpha < 3/2$ for $r \rightarrow 0$, i.e. the problem of stability.

We consider the three operators $K_n = \frac{\partial}{\partial x_n} + \beta \frac{x_n}{r^2}$ for $n = 1, 2, 3$ and let be $\phi \in C_0^\infty$. We have

$$0 \leq \sum_n (K_n \phi, K_n \phi) = \int \phi^* \left(-\nabla^2 - \frac{\beta}{r^2} + \frac{\beta^2}{r^2} \right) \phi d^3x$$

and for $\beta = 1/2$

$$0 \leq \int \phi^* \left(-\nabla^2 - \frac{1}{4r^2} \right) \phi \, d^3x .$$

It follows that

$$\begin{aligned} \left(\phi, \left(-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right) \phi \right) &= \left(\phi, \frac{\hbar^2}{2m} \left(-\nabla^2 - \frac{1}{4r^2} \right) \phi \right) + \left(\phi, \left(V(r) + \frac{\hbar^2}{2m} \frac{1}{4r^2} \right) \phi \right) \\ &\geq \min_r \left(V(r) + \frac{\hbar^2}{2m} \frac{1}{4r^2} \right) \end{aligned}$$

which is a finite energy.

One can show that the closure of H defined on $D_0(H) = C_0^\infty$ is the unique self-adjoint extension of H and, as under closure the lower bound is left unchanged, we have that the operator

$$-\frac{\hbar^2}{2m} \nabla^2 + V(r)$$

with $V(\infty) = 0$ and $|V(r)| < \text{const } r^{-3/2+\varepsilon}$ for $r \rightarrow 0$, see (5.1), is lower bounded.

If H is lower bounded as it is the rule in quantum mechanics, we have $H \geq bI$ and one can replace the spectral analysis of the unbounded self-adjoint operator H with the simpler analysis of the bounded self-adjoint positive operator $(H + cI)^{-1}$ where $c > -b$ and thus $H + cI > c_1 I > 0$. In fact we have

$$(H + cI)^{-1} = \int_0^{1/c_1} \lambda \, dE(\lambda)$$

from which

$$H = \int_0^{1/c_1} \frac{1 - c\lambda}{\lambda} \, dE(\lambda).$$

References

- [1] E.B. Davies *Spectral theory of differential operators*, Cambridge University Press, Cambridge (1995)

Chapter 6

Description of composite systems

We recall that if A and B are compatible observables we can write for any $|\psi\rangle \in \mathcal{H}$

$$|\psi\rangle = \sum_{abm} c_{ab}^m |a, b, m\rangle$$

where $|a, b, m\rangle$ are eigenstates of A to the value a and of B to the value b , and m takes into account the possible residual degeneracy.

$$\sum_m c_{ab}^m |a, b, m\rangle$$

is an eigenstate of A to the value a and eigenstate of B to the value b . Thus we can rewrite

$$|\psi\rangle = \sum_{ab} c_{ab} |a, b, (\psi)\rangle$$

where $|a, b, (\psi)\rangle$ are eigenstates of A and B relative to different pairs of values a, b .

Repeating the argument of Chapter 2 on the positivity of probabilities, performed for an observable we have that

$$\text{Prob}_{AB}^{(ab)}(\psi) = |c_{ab}|^2 = |\langle a, b, (\psi) | \psi \rangle|^2 = \langle \psi | P_{ab}^{A,B} | \psi \rangle \quad (6.1)$$

where P_{ab}^{AB} is the projector on the eigensubspace relative to the eigenvalue a of A and b of B . We come now to composite systems. As it happens in classical

mechanics, frequently one is faced with a system which is the composition of two subsystem.

Let system S be described by the Hilbert space $\mathcal{H}_S$ and system M by the Hilbert space $\mathcal{H}_M$.

We want to justify the following principle: The Hilbert space which describes the composite system $S + M$ is given by the tensor product space $\mathcal{H}_S \otimes \mathcal{H}_M$.

Let us consider first the case in which the two system are prepared independently and let A be an observable of S and B an observable of M . We have when system S is described by the vector $|S_1\rangle$ and system M by the vector $|M_1\rangle$

$$\text{Prob}_{AB}^{(ab)}(S_1 M_1) = \text{Prob}_A^{(a)}(S_1) \text{Prob}_B^{(b)}(M_1) = \langle S_1 | P_a^A | S_1 \rangle \langle M_1 | P_b^B | M_1 \rangle \quad (6.2)$$

due to the independence of the two systems.

The tensor product of the two Hilbert spaces is defined considering the pairs of vectors, one of $\mathcal{H}_S$ and the other of $\mathcal{H}_M$ and their linear combinations. The scalar product of two pairs of vectors is defined as the product of the two scalar products in $\mathcal{H}_S$ and $\mathcal{H}_M$

$$(\langle S_1 | \langle M_1 |)(|S_2\rangle |M_2\rangle) = \langle S_1 | S_2 \rangle \langle M_1 | M_2 \rangle$$

which is extended by linearity to all vectors. The product of the null vector of $\mathcal{H}_S$ and of any vector of $\mathcal{H}_M$ (and vice-versa) is defined as the null vector in the tensor product space.

One verifies immediately that such a space is a pre-Hilbert space. Its completion is the Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_M$, tensor product of $\mathcal{H}_S$ and $\mathcal{H}_M$.

One defines the action of the operators on $\mathcal{H}_S$, on the tensor product as follows

$$A(|S_1\rangle |M_1\rangle) = A|S_1\rangle |M_1\rangle$$

and the same for the operators B which act on $\mathcal{H}_M$. Then we have

$$[A, B] = 0$$

and thus two observables A and B relative to distinct subsystems are always compatible

$$A|a, m\rangle |b, n\rangle = a|a, m\rangle |b, n\rangle$$

$$B|a, m\rangle|b, n\rangle = b|a, m\rangle|b, n\rangle .$$

As $|a, m\rangle$ is a complete set in $\mathcal{H}_S$ and $|b, n\rangle$ a complete set in $\mathcal{H}_M$ we have that

$$|a, m\rangle|b, n\rangle$$

is a complete set in $\mathcal{H}_S \otimes \mathcal{H}_M$. We can write the result (6.2)

$$\text{Prob}_{AB}^{(ab)}(S_1 M_1) = \text{Prob}_A^{(a)}(S_1) \text{Prob}_B^{(b)}(M_1) = \langle S_1 | \langle M_1 | P_a^A P_b^B | S_1 \rangle | M_1 \rangle$$

which is consistent with the result (6.1).

For the wave functions i.e. when a basis is provided by the coordinates, if q_k is a complete set of compatible observables in $\mathcal{H}_S$ and Q_k a complete set of compatible observables in $\mathcal{H}_M$ we have that $|q\rangle|Q\rangle$ is a complete set vectors in the tensor product space and

$$\langle q | \langle Q | | S_1 \rangle | M_1 \rangle = \langle q | S_1 \rangle \langle Q | M_1 \rangle$$

i.e. the wave function of $|S_1\rangle|M_1\rangle$ is the product of the wave functions.

This happens only for factorized states. The wave function of a generic state will be a generic function of q and Q i.e. $\Psi(q, Q)$.

Chapter 7

The density operator

7.1 Introduction

Often we are in a situation in which the preparing apparatus does not produce a well defined quantum state but a statistical distribution of states.

There exists an analogue situation in classical physics when we describe the state of the system by a statistical ensemble, i.e. by means of a density function $\rho(q, p, t)$ in phase space and the mean values of the dynamical variables are given by

$$\bar{F} = \int F(q, p) \rho(q, p, t) dq dp / \int \rho(q, p, t) dq dp .$$

Sometime F may depend explicitly on time and in general one takes $\int \rho dq dp = 1$. We recall Liouville theorem in classical mechanics stating that in phase space the volume is conserved under the evolution due to Hamilton equations. In fact with $x_1 \dots x_n = q_1 \dots q_n$ and $x_{n+1} \dots x_{2n} = p_1 \dots p_n$ we have

$$\sum_j \frac{\partial \dot{x}_j}{\partial x_j} = 0 .$$

Thus with $x_j(t, x^0)$

$$V(t) = \int dx_1 \wedge \dots \wedge dx_{2n} = \int_{V^0} J dx_1^0 \wedge \dots \wedge dx_{2n}^0$$

where

$$J = \det \left(\frac{\partial x_j}{\partial x_k^0} \right) .$$

We have as a consequence of the Hamilton equations

$$\begin{aligned} \left. \frac{dV}{dt} \right|_{t=0} &= \int \frac{\partial \dot{x}_1}{\partial x_1} dx_1 \wedge \cdots \wedge dx_{2n} + \frac{\partial \dot{x}_2}{\partial x_2} dx_1 \wedge \cdots \wedge dx_{2n} \cdots \\ &= \int \frac{\partial \dot{x}_j}{\partial x_j} dx_1 \wedge \cdots \wedge dx_{2n} = 0 . \end{aligned}$$

Given a density function $\rho(q, p, t)$ in phase-space from the conservation of the number of points in phase space (determinism in both time directions) and the invariance of the volume we have

$$\frac{d\rho(q, p, t)}{dt} = 0 = \frac{\partial \rho(q, p, t)}{\partial t} + \frac{\partial \rho}{\partial q_j} \dot{q}_j + \frac{\partial \rho}{\partial p_j} \dot{p}_j \equiv \frac{\partial \rho(q, p, t)}{\partial t} + [\rho, H]_{PB}$$

i.e.

$$\frac{\partial \rho(q, p, t)}{\partial t} = -[\rho, H]_{PB} . \quad (7.1)$$

In the case of quantum mechanics we are in the situation in which we have more than one state ψ_j which occur with probability ρ_j with $\sum_j \rho_j = 1$. The mean value of the measures of the observable F on the ensemble of states is

$$\bar{F} = \sum_j \rho_j (\psi_j, F \psi_j) .$$

This can be written in more concise way as

$$\bar{F} = \text{Tr} (FW)$$

with W the density operator (sometime called also density matrix)

$$W = \sum_j \rho_j \psi_j \circ \psi_j , \quad \sum_j \rho_j = 1 , \quad \rho_j > 0 .$$

W is a bounded hermitean positive operator. The trace on a Hilbert space is defined by

$$\text{Tr} (A) = \sum_k (\zeta_k, A \zeta_k)$$

with ζ_k a complete orthonormal set in Hilbert space. For operators of the type WA with A bounded one can show that the trace $\text{Tr}(WA)$ is independent of the choice of the orthonormal set in Hilbert space. This because being W hermitean and positive we can write $W = VV$ with V hermitean and we have $\text{Tr} VV = 1$.

Thus VV is Hilbert-Schmidt and thus also WA is Hilbert-Schmidt for which the trace is independent of the used basis [1],[2].

We have

$$\begin{aligned}\text{Tr}(FW) &= \sum_k \sum_j (\zeta_k, F\psi_j) \rho_j(\psi_j, \zeta_k) = \sum_j \sum_k \rho_j(\psi_j, \zeta_k) (\zeta_k, F\psi_j) \\ &= \sum_j \rho_j(\psi_j, F\psi_j) .\end{aligned}$$

Notice that

$$\text{Tr}(W) = \sum_j \rho_j = 1 .$$

7.2 The time evolution of the density operator

The individual states evolve as

$$\psi_k(t) = e^{-iHt/\hbar} \psi_k(0) = U(t) \psi_k(0)$$

from which

$$W(t) = U(t)W(0)U^\dagger(t) .$$

Taking the derivative w.r.t. time we have in the Schrödinger picture

$$\frac{\partial}{\partial t} W(t) = -\frac{1}{i\hbar} [W(t), H(t)]$$

which is to be compared with the classical evolution equation (7.1). We have the following correspondence between classical and quantum mechanics.

$$\begin{aligned}\int dq dp &\leftrightarrow \text{Tr} \\ \rho(q, p, t) &\leftrightarrow W(t) \\ [,]_{PP} &\leftrightarrow \frac{1}{i\hbar} [,] .\end{aligned}$$

Pure states are represented by

$$W = \psi \circ \psi .$$

All other density operators are obtained by convex combination of pure states. Pure states are extremal in the ensemble of the convex set of density operators. If the state is pure we have

$$W^2 = W .$$

This is necessary but also sufficient for W to represent a pure state. In fact as $\text{Tr} W = 1$ we cannot have $W = 0$; thus there exists a ϕ such that

$$W\phi \neq 0 .$$

From $W^2 = W$ we have

$$WW\phi = W\phi \equiv \phi_1 \neq 0$$

and we choose $(\phi_1, \phi_1) = 1$. Let us construct an orthonormal basis in Hilbert space $\phi_1, \phi_2, \dots$. Being $W^+ = W$ we have

$$1 = \text{Tr}(W) = (\phi_1, W\phi_1) + (\phi_2, W\phi_2) + (\phi_3, W\phi_3) + \dots = 1 + (W\phi_2, W\phi_2) + (W\phi_3, W\phi_3) + \dots$$

and thus it follows that $W\phi_k = 0$ for $k \geq 2$. It follows that $W - \phi_1 \circ \phi_1 = 0$ as it is seen by applying such an operator to the basis ϕ_k .

By performing a sufficient number of measures it is possible to distinguish a statistical ensemble from pure states. Let us consider in fact the observables $Z = \zeta \circ \zeta$. We have

$$\bar{Z} = \text{Tr}(ZW) = (\zeta, W\zeta).$$

and we know that due to the arbitrariness of ζ the operator W is known and we can check the validity of $W = W^2$.

Let us consider now a composite system. A state vector ψ belonging to $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_M$ can always be written as

$$\psi = \sum_j \phi_j \xi_j$$

with $\phi_j \in \mathcal{H}_S$, $(\phi_j, \phi_j) = 1$ and $\xi_j \in \mathcal{H}_M$, $(\xi_j, \xi_k) = 0$ for $j \neq k$. Let us pose $(\xi_j, \xi_j) = \rho_j$. We have

$$1 = (\psi, \psi) = \sum_j (\xi_j, \xi_j) = \sum_j \rho_j .$$

If we measure an observable F relative to the subsystem S , as F does not operate on $\mathcal{H}_M$ we have

$$\bar{F} = (\psi, F\psi) = \sum_j (\xi_j, \xi_j)(\phi_j, F\phi_j) = \sum_j \rho_j(\phi_j, F\phi_j) = \text{Tr}_{\mathcal{H}_S}(WF)$$

where $W = \sum_j \rho_j \phi_j \circ \phi_j$ is a density operator which acts on $\mathcal{H}_S$ and the trace is performed on the Hilbert space $\mathcal{H}_S$.

The Hamiltonian of the system $S + M$ is given by

$$H = H_S + H_M + H_I$$

where H_S is the Hamiltonian of S and as such acts on $\mathcal{H}_S$, H_M is the Hamiltonian of M and as such acts on $\mathcal{H}_M$ while H_I is the interaction Hamiltonian which acts on $\mathcal{H}$.

If the initial state $\psi(0)$ is such that for $t > 0$ the interaction is negligible (e.g. in a decay process)

$$\psi(t) = e^{-i(H_S+H_M+H_I)t/\hbar}\psi(0) = e^{-i(H_S+H_M)t/\hbar}\psi(0)$$

and we have for $t > 0$

$$\begin{aligned} \bar{F}(t) &= (\psi(t), F\psi(t)) = (e^{-i(H_S+H_M)t/\hbar}\psi(0), Fe^{-i(H_S+H_M)t/\hbar}\psi(0)) = \\ &= (e^{-iH_S t/\hbar}e^{-iH_M t/\hbar}\psi(0), Fe^{-iH_S t/\hbar}e^{-iH_M t/\hbar}\psi(0)) = \\ &= \sum_j (e^{-iH_M t/\hbar}\xi_j(0), e^{-iH_M t/\hbar}\xi_j(0))(e^{-iH_S t/\hbar}\phi_j(0), Fe^{-iH_S t/\hbar}\phi_j(0)) = \\ &= \sum_j (\xi_j(0), \xi_j(0))(e^{-iH_S t/\hbar}\phi_j(0), Fe^{-iH_S t/\hbar}\phi_j(0)) = \\ &= \sum_j \rho_j(\phi_j(t), F\phi_j(t)) = \text{Tr}_{\mathcal{H}_S}(FW(t)) \end{aligned}$$

where $W(t)$ is the density operator evolved according to the Hamiltonian H_S .

References

- [1] N. Dunford and J. Schwarz *Linear operators*, Vol.II Chap.XI Interscience Publishers, New York, London (1963)
- [2] W. Thirring *Quantum mechanics of atoms and molecules*, Chap.2, Springer Verlag, (1979)

Chapter 8

Measure in quantum mechanics

We recall the results obtained in Chapter 2: Given a state ψ if we perform the measure of the observable F on ψ the probability to obtain the value f_i (of the spectrum) is given by $|c_i|^2$ being c_i the coefficient which appears in the expansion of ψ in eigenstates F i.e. $\psi = \sum_i c_i \phi_i$. Using the expansion in standard eigenstates $\psi = \sum_i c_i^{(n)} \phi_i^{(n)}$ such a probability is given by $\sum_n |c_i^{(n)}|^2$.

We justified such a formula generalizing Born hypothesis. A fundamental question which arises is: What happens after the measure, i.e. which is the vector that describes the system after the measure if we have obtained as the result of the measure the value f_i ?

A first classification of the measuring processes is the distinction between repeatable and non repeatable measures. We shall say that the measure of a certain quantity repeatable if repeating the measure of the same quantity on the system immediately after the first measure we obtain with certainty the result previously obtained.

An example of a repeatable measure is the Stern-Gerlach experiment. An example of a non repeatable measure is the measure of the energy of a particle through ionization.

Whether an apparatus effects a repeatable measure can be experimentally ascertained.

If the value for which we select the system is a non degenerate value of the spectrum and the measure is repeatable we conclude that the selected system

after the measure is in the eigenstate of F relative to the measured value f_i of the spectrum.

Thus in the case of repeatable measure in which we select the system if we measure a value f_i of the spectrum which is non degenerate, the measure operation coincides with that of the preparation of the state.

This can be expressed by saying that in such a measuring process the evolution $\psi \rightarrow \phi_i$ has taken place. We can describe such evolution by $\psi \rightarrow \phi_i = P_i \psi$ being P_i the projector $\phi_i \circ \phi_i$. Notice that such a linear transformation does not conserve the norm and as such it is not unitary. We recall also from Chap.4 that it is not possible to perform a further transformation which brings back the norm to 1 without losing linearity. Notice also that the probability to obtain f_i is given by $(\psi, P_i \psi)$.

More subtle is the problem when the measured value f_i is degenerate. In fact in this case it is not enough to say that the measure is repeatable to know the state of the system after the measure.

We shall introduce the notion of strongly repeatable measure or better of an apparatus which performs strongly repeatable measurements, characterized by the following property: If the initial state is an eigenstate of F , then the measure leaves such state unchanged. Also this is a property experimentally ascertainable. If this happens we shall say that our experimental apparatus performs a strongly repeatable measurement.

We do not want to give up linearity in quantum mechanics and thus we shall assume that the transition which occurs during the measure is given by a linear operator. Given an arbitrary initial state we denote by A_i the linear transformation which corresponds to the selection of the final states for which the result of the measure of F gave f_i ; these will be eigenstates of F to the value f_i . Example: The Stern and Gerlach experiment in which one keeps open just one slit.

On the eigenstates of ϕ of F to the value f_i , A_i acts as follows (strongly repeatable measurement)

$$A_i \phi = \alpha(\phi) \phi .$$

If we do not want to give up the linearity of the transformation we have $\alpha(\phi) = \alpha$ independent of ϕ and thus A_i/α is the identity operator on this subspace. On

the eigenvectors of F to a different value f_j of the spectrum we have

$$A_i \phi_j = 0$$

otherwise a state which gives with certainty the value f_j would pass the selection test A_i which is contradictory with the probability formula for the measure of f_i .

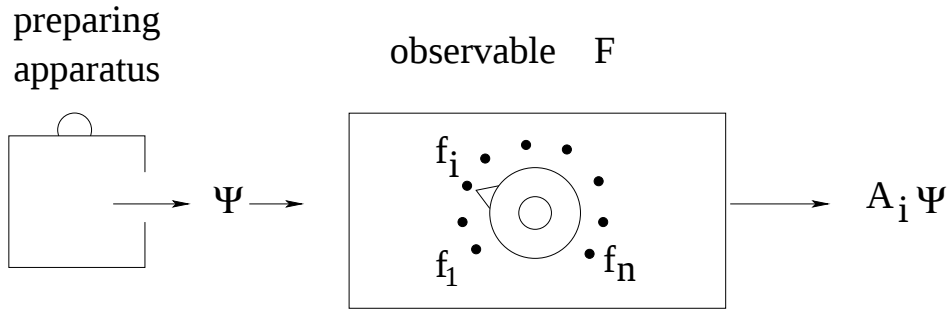


Figure 8.1: Generalized Stern-Gerlach apparatus: the measuring apparatus passes only those systems for which the value f_i of the spectrum of F has been measured.

We conclude that A_i apart for a factor is the projection operator P_i on the eigensubspace of F to the eigenvalue f_i . Such a projection property of the strongly repeatable measures is called for obvious reasons, the property of least disturbance. Notice again that the probability to find the value f_i is given by $\text{Prob}(f_i) = (\psi, P_i \psi)$, with $(\psi, \psi) = 1$.

If our system is described by a density operator it is easy to find how this changes due to a measuring process. If we select the systems for which the measuring apparatus has measured the value f_i , as any state ψ is changed into $P_i \psi / \|P_i \psi\|$ (strongly repeatable measure) and this happens with probability $(P_i \psi, P_i \psi)$ we have that the density operator after the measure is given by $P_i W P_i$. Notice that such a density operator is no longer normalized i.e. in general we have $\text{Tr } P_i W P_i < 1$. This is due to the fact that a certain number of systems are lost. We can normalize by dividing by its trace i.e. the new normalized density operator is given by

$$W' = \frac{P_i W P_i}{\text{Tr } (W P_i)}.$$

We recall that the probability to find f_i in the measure of F on W is given by $\sum_j \rho_j(\psi_j, P_i \psi_j)$ i.e. by $\text{Tr}(WP_i)$.

We want now, given two observables F and G , compute the composite probability to find in two immediately successive measurements first f_i and then g_j . Such a probability is given by

$$\text{Prob}(f_i, g_j) = \text{Prob}(f_i) \text{Prob}(g_j|f_i) = \text{tr}(WP_i)\text{tr}(W'Q_j) = \text{tr}(P_iWP_iQ_j).$$

We could proceed with a sequence of many measures, but we shall limit ourselves to two measures.

On the other hand the probability to find on W first g_j and thereafter f_i is given by

$$\text{Prob}(g_j, f_i) = \text{tr}(Q_jWQ_jP_i).$$

In general these two probabilities are different, contrary to what we know to happen experimentally in the classical case.

It is easy to see that if P_i commutes with Q_j then the two composite probabilities are equal: In fact in this case $P_iQ_jP_i = P_i^2Q_j = P_iQ_j = Q_jP_iQ_j$.

Let us suppose now to find experimentally that the two above mentioned composite probabilities are equal, whatever the state W and for all indices i and j . Given the arbitrariness of the state W , (think e.g. to the pure states) we have that $P_iQ_jP_i = Q_jP_iQ_j$ for all choices of indices i and j . It is easy then to see that this implies $[P_i, Q_j] = 0$, for all pairs i, j . In fact

$$P_i = P_i^2 = \sum_k P_iQ_kP_i = \sum_k Q_kP_iQ_k$$

from which we derive

$$P_iQ_j = Q_jP_iQ_j = Q_jP_i$$

which is what we wanted to prove. If we now consider the spectral representations of the operators

$$F = \sum_i f_i P_i \qquad G = \sum_j g_j Q_j$$

we have that $[F, G] = 0$. Vice-versa if F and G commute we have that there exist a complete set ϕ_{ijn} of eigenstates common to F and G i.e.

$$F\phi_{ijn} = f_i\phi_{ijn} \qquad G\phi_{ijn} = g_j\phi_{ijn}$$

from which

$$P_i = \sum_{k,n} \phi_{ikn} \circ \phi_{ikn} \quad Q_j = \sum_{h,m} \phi_{hjm} \circ \phi_{hjm}$$

and

$$P_i Q_j = \sum_n \phi_{ijn} \circ \phi_{ijn} = Q_j P_i.$$

In conclusion: The temporal commutativity of the measure of F and G at the statistical level implies the mathematical commutativity of the operators which represent the observables F and G and vice-versa.

References

- [1] K. Gottfried *Quantum Mechanics*, W. A. Benjamin, New York (1966)
- [2] E.B. Davies *Quantum mechanics of open systems*, Academic Press, London (1976)
- [3] J.S. Bell *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge U.K. (1993)
- [4] J.A. Wheeler and W.H.Zurek *Quantum theory of measurement*, Princeton University Press, Princeton (1983)

Chapter 9

Symmetry transformations

9.1 Space translations

As an introduction we want to find the generator of translations. Let us consider a system described by the vector ψ and let

$$\psi_\varepsilon = (1 - i\varepsilon\hat{t}_k)\psi$$

be the state vector which describes the same system translated by ε in the positive direction along the k axis. We must have

$$(\psi_\varepsilon, \hat{q}_m \psi_\varepsilon) = (\psi, \hat{q}_m \psi) + \varepsilon \delta_{km}$$

and thus

$$[\hat{q}_m, \hat{t}_k] = i\delta_{mk} .$$

Similarly

$$(\psi_\varepsilon, \hat{p}_m \psi_\varepsilon) = (\psi, \hat{p}_m \psi)$$

that is

$$[\hat{p}_m, \hat{t}_k] = 0.$$

The first tells us that $\hat{t}_k = \hat{p}_k/\hbar + f_k(\hat{\mathbf{q}})$, where $f_k(\mathbf{q})$ are three functions of the coordinates, while the second tells us that f_k are constants. Thus we have proved that

$$\hat{t}_k = \hat{p}_k/\hbar + c_k$$

where the c_k , if we want $(1 - i\varepsilon\hat{t}_k)$ to conserve the norm of vectors, must be real numbers. The finite translation is given by

$$e^{-i\hat{p}_k b/\hbar} e^{-ic_k b}.$$

Being the group of translations abelian translations are easily composed.

One could even ask whether we cannot define the momentum as the generator of translations and similarly angular momentum as the generator of rotations. The advantage is to allow representations of the angular momentum more general than the orbital angular momentum i.e. the spin.

To carry out such a program is necessary to develop some general considerations on symmetry transformations.

9.2 Symmetry transformations

Let us consider for concreteness the orthogonal transformation $\mathbf{q}' = \gamma(\mathbf{q})$ with γ a proper orthogonal matrix, $\det \gamma = 1$.

We shall assume the active viewpoint: Rotating the system (with respect to the “fixed stars” or better with respect to the cosmic microwave radiation) the physical points move from the positions of coordinates $\mathbf{q}$ to those of coordinates $\mathbf{q}' = \gamma(\mathbf{q})$.

For microscopic systems this is obtained by rotating by γ the preparing apparatus and the rotated system will be described by a vector $T\psi$.

In presence of invariance under rotations (i.e. in absence of external fields) the invariance of the transition probability tells us

$$\frac{|(\phi, \psi)|^2}{(\phi, \phi)(\psi, \psi)} = \frac{|(T\phi, T\psi)|^2}{(T\phi, T\phi)(T\psi, T\psi)} \quad (9.1)$$

Notice that (9.1) is an experimentally verifiable relation.

A set of transformation T which admit inverse and for which (9.1) holds, form a group and such a group is called a symmetry group.

The transformation T can always be written such that it conserves the norm as up to now T is a general, not necessarily linear, transformation; i.e. we shall set

$$(T\psi, T\psi) = (\psi, \psi). \quad (9.2)$$

Clearly a change of type

$$T \rightarrow T' \quad \text{defined by} \quad T'\psi = \exp(i\alpha(\psi))T\psi \quad (9.3)$$

leaves invariant the relations (9.1) and (9.2) and thus we can exploit such arbitrariness to reduce the operator T to a more conventional form. Wigner [1] [2] shows that it is always possible to choose the phases in (9.3) in such a way that the operator T becomes either linear or anti-linear. In the linear case (9.2) tells us that the linear operator $T = L$ is isometric i.e. $L^+L = 1$ and as T has inverse we have also $LL^+ = 1$ i.e. L is unitary.

Let us examine now the anti-linear case. The definition of anti-linear operator is

$$A(\alpha\psi + \beta\phi) = \alpha^*A\psi + \beta^*A\phi.$$

Given $(A\psi, A\psi) = (\psi, \psi)$ from the definition of anti-linear operator we have

$$\begin{aligned} (A(\alpha\psi + \beta\phi), A(\alpha\psi + \beta\phi)) &= |\alpha|^2(\psi, \psi) + \beta^*\alpha(\phi, \psi) + \alpha^*\beta(\psi, \phi) + |\beta|^2(\phi, \phi) \\ &= |\alpha|^2(A\psi, A\psi) + \beta^*\alpha(A\psi, A\phi) + \alpha^*\beta(A\phi, A\psi) + |\beta|^2(A\phi, A\phi) \end{aligned}$$

and due to the arbitrariness of α and β we have $(A\psi, A\phi) = (\phi, \psi)$ which defines an anti-isometric operator. The complex number $(\psi, A\phi)^* = (A\phi, \psi)$ depends linearly on ϕ and thus can be written due to Riesz theorem as (ζ, ϕ) i.e.

$$(\psi, A\phi) = (\phi, \zeta)$$

with ζ depending anti-linearly on ψ and thus we can introduce the anti-linear operator A^+ which we shall call the adjoint of A defined by

$$A^+\psi = \zeta, \quad (\psi, A\phi) = (\phi, A^+\psi) = (A^+\psi, \phi)^*.$$

The invariance of the norm of ψ tells us that such operator A is anti-isometric i.e. $A^+A = 1$ while again the invertibility property tell us that also that $AA^+ = 1$ and A is called anti-unitary. Such arguments hold for all symmetry operations. We prove now [1] that for all symmetry operations which projectively commute with the time evolution and transform energy eigenstates in energy eigenstates with the same value of the energy the only possibility, consistent with the superposition principle is the unitary case. Let us consider in fact the superposition of two

eigenstates of the energy ψ_1 and ψ_2 relative to different values of the energy E_1 and E_2 and suppose that the symmetry transformation is anti-unitary. The state $\alpha_1\psi_1 + \alpha_2\psi_2$ at time 0 evolves at time t into the state

$$\alpha_1 \exp(-iE_1 t/\hbar)\psi_1 + \alpha_2 \exp(-iE_2/\hbar)\psi_2. \quad (9.4)$$

Instead the transformed state at $t = 0$ i.e. $\alpha_1^* A\psi_1 + \alpha_2^* A\psi_2$ evolves at time t into the state

$$\alpha_1^* \exp(-iE_1/\hbar)A\psi_1 + \alpha_2^* \exp(-iE_2/\hbar)A\psi_2. \quad (9.5)$$

If we now transform according to the operator A the state (9.4) we must find the state (9.5) and thus the vector

$$\alpha_1^* \exp(iE_1/\hbar)A\psi_1 + \alpha_2^* \exp(iE_2/\hbar)A\psi_2$$

may differ at most by a phase factor from

$$\alpha_1^* \exp(-iE_1/\hbar)A\psi_1 + \alpha_2^* \exp(-iE_2/\hbar)A\psi_2.$$

But being the two states $A\psi_1$ and $A\psi_2$ orthogonal and $E_1 \neq E_2$ this cannot be valid for all t . Thus the hypothesis of a anti-unitary transformation gives rise to a contradiction. In particular the two statements $[A, H] = 0$ and $\exp(-iHt/\hbar)A = A \exp(-iHt/\hbar)$ with A anti-unitary are in contradiction. On the contrary the following two relations are consistent $[A, H] = 0$ and $\exp(iHt/\hbar)A = A \exp(-iHt/\hbar)$, with A anti-linear; actually from the first we obtain the second e.g. by expanding in series the exponential and recalling that A anti-commutes with i . If there exists an A anti-linear with the property of commuting with H then we say that A is an operation of inversion of time and that the system described by H is invariant under time reversal.

Let us consider a group of symmetry transformations represented by unitary operators and let γ_1 and γ_2 be represented by the unitary operators $U(\gamma_1)$ and $U(\gamma_2)$. To the product transformation $\gamma_2\gamma_1$ there corresponds the unitary operator $U(\gamma_2\gamma_1)$. But the sequence of the two transformations γ_1, γ_2 is physically equivalent to the transformation $\gamma_2\gamma_1$ and thus we must have

$$U(\gamma_2\gamma_1)\psi = \alpha(\gamma_2, \gamma_1, \psi)U(\gamma_2)U(\gamma_1)\psi$$

where α is a phase factor which in principle can depend both on γ_1, γ_2 and on the state vector ψ .

A simple reasoning shows that α cannot depend on the vector ψ . In fact let us consider two unitary operators U and V such that for any vector ψ holds $U\psi = \alpha V\psi$ with α in general dependent on ψ . Defined $K = V^+U$ and given two linearly independent vectors ψ_1 and ψ_2 we have $K\psi_1 = \alpha_1\psi_1$ and $K\psi_2 = \alpha_2\psi_2$ and also

$$\begin{aligned} K(a_1\psi_1 + a_2\psi_2) &= a_1K\psi_1 + a_2K\psi_2 = a_1\alpha_1\psi_1 + a_2\alpha_2\psi_2 \\ &= \alpha_3(a_1\psi_1 + a_2\psi_2) = a_1\alpha_3\psi_1 + a_2\alpha_3\psi_2 \end{aligned}$$

and being ψ_1 and ψ_2 independent, we must have $\alpha_3 = \alpha_1$ and $\alpha_3 = \alpha_2$, i.e. $\alpha_1 = \alpha_2$ independent of ψ_j . Thus

$$U(\gamma_2\gamma_1) = \alpha(\gamma_2, \gamma_1)U(\gamma_2)U(\gamma_1). \quad (9.6)$$

A correspondence $\gamma \rightarrow U(\gamma)$ which satisfies (9.6) is called a representation up to a phase, or a projective representation, of the symmetry group.

The following theorem by Bargmann holds [3]: Given a projective representation of a compact group i.e. a representation which satisfies (9.6), continuous in a neighborhood of the identity it is possible to perform a phase transformation on the U i.e. $U \rightarrow \omega(U)U$ with ω phase factor, such that in a neighborhood of the identity it remains continuous and in such a neighborhood becomes a true representation i.e. (9.6) holds with $\alpha \equiv 1$. The same cannot be said for the representation in the large.

In fact we can assure that (9.6) holds with $\alpha \equiv 1$ for any pair of transformations only if the group is simply connected. As we shall see in Chapter 10, $SO(3)$ is not simply connected and in general we cannot redefine the phases in such a way as to have a continuous representation with $\alpha \equiv 1$ in (9.6) in the large.

9.3 Rotations

Let us specify our considerations to the group of rotations i.e. $SO(3)$. The main feature which distinguishes this group from the group of translation is its non commutativity.

Given a real antisymmetric transformation in three dimensions $\alpha = -\alpha^T$, let us consider e^α . We have $(e^\alpha)^T = e^{-\alpha} = (e^\alpha)^{-1}$, i.e. e^α is an orthogonal transformation and $e^{\phi\alpha}$ is a one parameter group of orthogonal transformations connected with the identity. We recall that in three dimensions all proper rotations i.e. the elements of $SO(3)$ are rotations around an axis.

Given two infinitesimal rotations characterized by the antisymmetric transformation α and β we want to compute the “commutator” of the transformations generated by α and β i.e. $\exp(-\varepsilon\beta)\exp(-\varepsilon\alpha)\exp(\varepsilon\beta)\exp(\varepsilon\alpha)$ keeping up to the second order terms. Using twice the Baker-Campbell-Hausdorff formula $\exp(A)\exp(B) = \exp(A + B + [A, B]/2) + O_3$ we obtain

$$e^{-\varepsilon\beta}e^{-\varepsilon\alpha}e^{\varepsilon\beta}e^{\varepsilon\alpha} = e^{-\varepsilon\beta-\varepsilon\alpha+\varepsilon^2[\beta,\alpha]/2+O(\varepsilon^3)}e^{\varepsilon\beta+\varepsilon\alpha+\varepsilon^2[\beta,\alpha]/2+O(\varepsilon^3)} = e^{\varepsilon^2[\beta,\alpha]+O(\varepsilon^3)} \quad (9.7)$$

(notice that also $[\beta, \alpha]$ is an antisymmetric transformation). According to Wigner’s theorem, given a rotation, it is possible to represent it in Hilbert space by means of a unitary transformation and according to Bargmann theorem the $U(\gamma)$ can be chosen as to form a true representation in a neighborhood of the identity i.e. such that

$$U(\gamma_2)U(\gamma_1) = U(\gamma_2\gamma_1). \quad (9.8)$$

According to Stone theorem the one parameter group of unitary operators $U(e^{\phi\alpha})$ can be written in the form

$$U(e^{\phi\alpha}) = e^{-i\phi r(\alpha)/\hbar} \quad (9.9)$$

with $r(\alpha)$ self-adjoint operator and the same for β . Using again the Baker-Campbell-Hausdorff formula we obtain

$$U(e^{-\varepsilon\beta})U(e^{-\varepsilon\alpha})U(e^{\varepsilon\beta})U(e^{\varepsilon\alpha}) = e^{\varepsilon^2[-ir(\beta)/\hbar, -ir(\alpha)/\hbar]+O(\varepsilon^3)}$$

which according to (9.7) and to (9.8) must be equal to

$$e^{-i\varepsilon^2 r[\beta, \alpha]/\hbar + O(\varepsilon^3)}$$

i.e.

$$[r(\alpha), r(\beta)] = i\hbar r([\alpha, \beta]). \quad (9.10)$$

Let us specify now the transformations α and β as generators of rotations around the coordinate axes.

$$\alpha_1 = A_1, \quad \alpha_2 = A_2, \quad \alpha_3 = A_3$$

with A_1, A_2, A_3 given by

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

E.g. α_3 gives rise to the infinitesimal transformation

$$\begin{cases} q'_1 = q_1 - \epsilon_3 q_2 \\ q'_2 = q_2 + \epsilon_3 q_1 \\ q'_3 = q_3 \end{cases}$$

The three matrices A satisfy the commutation relations

$$A_i A_j - A_j A_i = \epsilon_{ijk} A_k \quad \text{e.g.} \quad A_1 A_2 - A_2 A_1 = A_3. \quad (9.11)$$

Setting now

$$r(A_j) = M_j$$

and using (9.10) and (9.11) we obtain

$$[M_j, M_k] = i\hbar \epsilon_{jkl} M_l \quad (9.12)$$

which are identical to the commutation rules of the orbital angular momentum which can be derived by the transcription of the classical angular momentum of a particle (coordinate representation).

However presently we reached (9.12) without the need to specify the nature of the vector we are transforming.

To start, let us consider the case in which the state is represented by the wave function $\psi(\mathbf{q})$. The simplest and most natural way to transform the wave function under rotations is obtained imposing the invariance in value of the wave function i.e.

$$\psi'(\mathbf{q}') = \psi'(\gamma_1(\mathbf{q})) = \psi(\mathbf{q}) \quad \text{i.e.} \quad \psi'(\mathbf{q}) = \psi(\gamma_1^{-1}(\mathbf{q})) = U(\gamma_1)\psi(\mathbf{q}). \quad (9.13)$$

Under the sequence of two transformations γ_1 and then γ_2 we have

$$\begin{aligned} U(\gamma_2)U(\gamma_1)\psi(\mathbf{q}) &= U(\gamma_2)\psi(\gamma_1^{-1}(\mathbf{q})) = \psi(\gamma_1^{-1}\gamma_2^{-1}(\mathbf{q})) \\ &= \psi((\gamma_2\gamma_1)^{-1}(\mathbf{q})) = U(\gamma_2\gamma_1)\psi(\mathbf{q}) . \end{aligned}$$

It is immediate that transformation (9.13) is unitary insofar

$$\int \psi^*(\gamma^{-1}(\mathbf{q}))\phi(\gamma^{-1}(\mathbf{q})) d\mathbf{q} = \int \psi^*(\mathbf{q})\phi(\mathbf{q}) d\mathbf{q}$$

γ being a proper orthogonal transformation. Then we have $\psi'(\mathbf{q}) = \psi(\gamma_1^{-1}(\mathbf{q})) \equiv U(\gamma_1)\psi(\mathbf{q})$ and $U(\gamma_2\gamma_1) = U(\gamma_2)U(\gamma_1)$ without any additional phase. Thus the transformation in value of the wave functions realizes completely the program of obtaining for $SO(3)$ a true representation of the whole group. Let us see the explicit form of the generator of the rotations around the axis 3.

$$\begin{aligned} (1 - i\varepsilon r(A_3)/\hbar)\psi(\mathbf{q}) &= (1 - i\varepsilon M_3/\hbar)\psi(\mathbf{q}) = \\ &= \psi(\mathbf{q} - \varepsilon A_3(\mathbf{q})) = \psi(q_1 + \varepsilon q_2, q_2 - \varepsilon q_1, q_3) = \psi(\mathbf{q}) + \varepsilon(q_2 \frac{\partial}{\partial q_1} - q_1 \frac{\partial}{\partial q_2})\psi(\mathbf{q}) \\ &= \psi(\mathbf{q}) - \frac{i}{\hbar}\varepsilon(-i\hbar q_1 \frac{\partial}{\partial q_2} + i\hbar q_2 \frac{\partial}{\partial q_1}) \end{aligned}$$

i.e.

$$M_3 = q_1 p_2 - q_2 p_1$$

which is exactly the expression of the orbital angular momentum.

9.4 The euclidean group

We want now to give the complete treatment of the euclidean group [4].

The translations

$$q'_k = q_k + c_k$$

combined with the rotations

$$q'_k = \gamma_{kl} q_l$$

form obviously a group which is called the euclidean group.

If we denote by T_k the generators of translations and J_k those of the rotations ($J_k \equiv A_k$) we obtain easily the following commutation rules

$$[J_k, J_l] = \varepsilon_{klm} J_m$$

$$[J_k, T_l] = \varepsilon_{klm} T_m$$

$$[T_k, T_l] = 0 .$$

The last one tells us simply that the translations commute among themselves. In quantum mechanics the symmetry group “euclidean group” can be projectively represented by unitary transformations (Wigner theorem) whose generators are self-adjoint operators (Stone theorem). The group however is not compact and thus we cannot exploit Bargmann theorem to reach a true representation in a neighborhood of the identity. If we denote by $\hat{J}_k$ and $\hat{T}_k$ the operators which generate the rotations and the translations on the vectors in Hilbert space, due to the fact that in quantum mechanics what are essential are the rays and not the vectors we can only write

$$[\hat{J}_k, \hat{J}_l] = i\hbar(\varepsilon_{klm} \hat{J}_m + a_{kl})$$

$$[\hat{J}_k, \hat{T}_l] = i\hbar(\varepsilon_{klm} \hat{T}_m + b_{kl})$$

$$[\hat{T}_k, \hat{T}_l] = i\hbar c_{kl}$$

where due to the unitarity of the transformations a_{kl} , b_{kl} , c_{kl} have to be real. Notice that being the commutators antisymmetric we have that $a_{kl} = -a_{lk}$, $c_{kl} = -c_{lk}$ while the same cannot be said for the b_{kl} . Given the antisymmetry we can write

$$a_{kl} = \varepsilon_{klm} a_m; \quad c_{kl} = \varepsilon_{klm} c_m .$$

Jacobi identity

$$[\hat{J}_k, [\hat{T}_l, \hat{T}_m]] + \text{cyclic} = 0$$

furnishes

$$(\varepsilon_{mkr} \varepsilon_{lrs} + \varepsilon_{klr} \varepsilon_{mrs}) c_s = 0$$

i.e. $c_s = 0$. Decomposing b_{kl} in antisymmetric and symmetric part

$$b_{kl} = \varepsilon_{klm} b_m + b_{(kl)}$$

Jacobi identity

$$[\hat{J}_k, [\hat{J}_l, \hat{T}_m]] + \text{cyclic} = 0$$

furnishes

$$\varepsilon_{lmr}b_{(kr)} + \varepsilon_{mkr}b_{(rl)} - \varepsilon_{klr}b_{(mr)} = 0 . \quad (9.14)$$

Anti-symmetrizing (9.14) in lm we obtain

$$\varepsilon_{lmr}b_{(kr)} = 0; \quad \text{i.e.} \quad b_{(kr)} = 0 .$$

Thus we reached

$$[\hat{J}_k, \hat{J}_l] = i\hbar\varepsilon_{klm}(\hat{J}_m + a_m)$$

$$[\hat{J}_k, \hat{T}_l] = i\hbar\varepsilon_{klm}(\hat{T}_m + b_m)$$

$$[\hat{T}_k, \hat{T}_l] = 0 .$$

Defining $\tilde{J}_k = \hat{J}_k + a_k$, $\tilde{T}_k = \hat{T}_k + b_k$ we have

$$[\tilde{J}_k, \tilde{J}_l] = i\hbar\varepsilon_{klm}\tilde{J}_m \quad (9.15)$$

$$[\tilde{J}_k, \tilde{T}_l] = i\hbar\varepsilon_{klm}\tilde{T}_m \quad (9.16)$$

$$[\tilde{T}_k, \tilde{T}_l] = 0 . \quad (9.17)$$

The (9.15,9.16,9.17) are called a *canonical representation* of rotations and translations. Eq.(9.15) gives rise to the quantization of angular momentum as already discussed (and experimentally observed) while the (9.17) combined with

$$[\hat{q}_k, \hat{q}_l] = 0; \quad [\hat{q}_k, \hat{T}_l] = i\delta_{kl}$$

can be used as the general definition of momentum: $\hat{p}_k = \hbar\hat{T}_k$.

Summing up: We define momentum and angular momentum of a quantum system the *canonical representation* of the generators of the euclidean group.

References

- [1] E.P. Wigner *Group theory*, Academic Press, New York (1959)
- [2] S. Weinberg *The quantum theory of fields* I, Chap. 2, Cambridge University Press (1995)

- [3] V. Bargmann *On unitary ray representations of continuous groups*, Ann.of Math. **59** 1 (1952))
- [4] G. Sartori *Lezioni di meccanica quantistica*, Edizioni Libreria Cortina, Padova (1998)

Chapter 10

The spin

10.1 Realization of half-integer angular momenta

We look for a Hilbert space $\mathcal{H}$ on which three hermitean operators s_i , $i = 1, 2, 3$ are defined which satisfy the angular momentum commutation relations i.e.

$$s_j s_k - s_k s_j = i \sum_l \epsilon_{jkl} s_l$$

and such that also $\sum_j s_j^2 = 1/2(1/2 + 1)$. It is useful to set $\mathbf{s} = \frac{1}{2}\boldsymbol{\sigma}$ from which

$$\sigma_j \sigma_k - \sigma_k \sigma_j = 2 i \sum_l \epsilon_{jkl} \sigma_l . \quad (10.1)$$

Let χ_+ and χ_- be the eigenvectors of σ_3

$$\sigma_3 \chi_+ = \chi_+; \quad \sigma_3 \chi_- = -\chi_- .$$

Clearly we have $(\chi_-, \chi_+) = 0$ and being $\frac{1}{2}$ the maximum eigenvalue of s_3 , $(\sigma_1 + i\sigma_2)\chi_+ \equiv \sigma_+ \chi_+ = 0$. We shall define the phase of χ_- by means of

$$(\sigma_1 - i\sigma_2)\chi_+ \equiv \sigma_- \chi_+ = 2\chi_- .$$

From this it follows

$$(\chi_+, \sigma_3 \chi_+) = 1; \quad (\chi_+, \sigma_3 \chi_-) = 0$$

$$(\chi_-, \sigma_3 \chi_+) = 0; \quad (\chi_-, \sigma_3 \chi_-) = -1$$

while the matrix elements of σ_+ and σ_- are all zero except

$$(\chi_-, \sigma_- \chi_+) = 2, \quad (\chi_+, \sigma_+ \chi_-) = 2.$$

Thus we have the following two-dimensional representation

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

called Pauli matrices. Notice that $\sigma_j^2 = 1$ which tells us simply that every σ_k has eigenvalues ± 1 .

From the commutation relations of the σ_j and from $\sigma_j^2 = 1$ it follows also that σ_j obey a Clifford algebra

$$\{\sigma_j, \sigma_k\} = 2 \delta_{jk}$$

where $\{, \}$ denotes the anticommutator.

This is found e.g. by summing to the (10.1) multiplied on the right by σ_k , the same multiplied on the left by σ_k ($k \neq j$).

According to eq.(9.9) the infinitesimal rotation around an axis $\mathbf{n}$ of an angle ϵ is given by

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \left(1 - \frac{i}{\hbar} \mathbf{M} \cdot \mathbf{n} \epsilon\right) \begin{pmatrix} a \\ b \end{pmatrix} = \left(1 - \frac{i}{2} \boldsymbol{\sigma} \cdot \mathbf{n} \epsilon\right) \begin{pmatrix} a \\ b \end{pmatrix}$$

and the finite rotation by

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \exp\left(-\frac{i}{2} \boldsymbol{\sigma} \cdot \mathbf{n} \phi\right) \begin{pmatrix} a \\ b \end{pmatrix} = \left(\cos\left(\frac{\phi}{2}\right) - i \boldsymbol{\sigma} \cdot \mathbf{n} \sin\left(\frac{\phi}{2}\right)\right) \begin{pmatrix} a \\ b \end{pmatrix}$$

as $(\boldsymbol{\sigma} \cdot \mathbf{n})^2 = 1$. We see that for a rotation of 2π around any axis we have $\begin{pmatrix} a' \\ b' \end{pmatrix} = -\begin{pmatrix} a \\ b \end{pmatrix}$ i.e. the vector changes of sign but this does not contrast with the physical interpretation of the state vector.

The 2×2 matrices

$$U = \cos\left(\frac{\phi}{2}\right) - i \boldsymbol{\sigma} \cdot \mathbf{n} \sin\left(\frac{\phi}{2}\right) \quad (10.2)$$

are all the elements of the group $SU(2)$ i.e. the group of the unimodular (i.e. determinant 1) unitary transformations in two dimensions.

In fact introducing the identity matrix in two dimensions σ_0 we have that writing the most general two-dimensional matrix in the form $a\sigma_0 + \mathbf{b} \cdot \boldsymbol{\sigma}$ its determinant

is given by $a^2 - \mathbf{b}^2$. The inverse of the above written unimodular matrix is given by $a\sigma_0 - \mathbf{b} \cdot \boldsymbol{\sigma}$ as is immediately verified using Clifford algebra.

If we impose now unitarity i.e. that the inverse equals the adjoint we have that $a = a^*$ and $\mathbf{b} = -\mathbf{b}^*$. Thus the most general $SU(2)$ matrix is written in the form

$$a\sigma_0 + i\mathbf{b} \cdot \boldsymbol{\sigma} \quad (10.3)$$

with real a and $\mathbf{b}$ and $a^2 + \mathbf{b}^2 = 1$ i.e. in one to one correspondence with the point of the surface of a four-dimensional sphere of radius 1 which is a simply connected set.

We want now to prove that (10.2) is a projective representation of the group $SO(3)$. To this end it is necessary to show that given two elements of $SO(3)$, γ_1 and γ_2 , for the related elements of $SU(2)$, $U(\gamma_1)$ and $U(\gamma_2)$ we must have

$$U(\gamma_2\gamma_1) = \alpha(\gamma_2, \gamma_1)U(\gamma_2)U(\gamma_1) \quad (10.4)$$

with $\alpha(\gamma_2, \gamma_1)$ a phase factor. Given an element of $SO(3)$, i.e. the rotation around an axis $\mathbf{n}$ of an angle ϕ we consider the element of $SU(2)$ related according to formula (10.2) which we already saw to be determined up to a sign. Let us show that

$$U^+(\gamma)\boldsymbol{\sigma}U(\gamma) = \gamma(\boldsymbol{\sigma}) . \quad (10.5)$$

If we perform a rotation of an angle ϕ around the axis 3 we have

$$\begin{aligned} \left(\cos\left(\frac{\phi}{2}\right) + i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) \sigma_1 \left(\cos\left(\frac{\phi}{2}\right) - i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) &= \cos \phi \sigma_1 - \sin \phi \sigma_2 \\ \left(\cos\left(\frac{\phi}{2}\right) + i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) \sigma_2 \left(\cos\left(\frac{\phi}{2}\right) - i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) &= \sin \phi \sigma_1 + \cos \phi \sigma_2 \\ \left(\cos\left(\frac{\phi}{2}\right) + i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) \sigma_3 \left(\cos\left(\frac{\phi}{2}\right) - i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) &= \sigma_3 \end{aligned}$$

i.e. we found that

$$\left(\cos\left(\frac{\phi}{2}\right) + i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) \boldsymbol{\sigma} \left(\cos\left(\frac{\phi}{2}\right) - i\sigma_3 \sin\left(\frac{\phi}{2}\right) \right) = \gamma(\boldsymbol{\sigma})$$

It is not difficult to prove that (10.5) holds for the rotation of an angle ϕ around whatever axis $\mathbf{n}$.

In fact under the rotation around $\mathbf{n}$ of an angle ϕ we have

$$\begin{aligned}\mathbf{q} \rightarrow \mathbf{q}' &= \gamma(\mathbf{q}) = \\ &= \mathbf{n}(\mathbf{q} \cdot \mathbf{n}) + \cos \phi (\mathbf{q} - \mathbf{n}(\mathbf{q} \cdot \mathbf{n})) + \sin \phi \mathbf{n} \wedge \mathbf{q} .\end{aligned}$$

Using the relation $(\boldsymbol{\sigma} \cdot \mathbf{n}) \sigma_k (\boldsymbol{\sigma} \cdot \mathbf{n}) = -\sigma_k + 2 n_k (\boldsymbol{\sigma} \cdot \mathbf{n})$ which is easily derived from Clifford algebra we find

$$\left(\cos\left(\frac{\phi}{2}\right) + i \mathbf{n} \cdot \boldsymbol{\sigma} \sin\left(\frac{\phi}{2}\right) \right) \boldsymbol{\sigma} \left(\cos\left(\frac{\phi}{2}\right) - i \mathbf{n} \cdot \boldsymbol{\sigma} \sin\left(\frac{\phi}{2}\right) \right) = \gamma(\boldsymbol{\sigma}) .$$

Thus given two elements of $SO(3)$, γ_1 and γ_2 and considered their product $\gamma_2\gamma_1$ we have

$$\begin{aligned}U^+(\gamma_1)U^+(\gamma_2) \boldsymbol{\sigma} U(\gamma_2)U(\gamma_1) &= U^+(\gamma_1) \gamma_2 (\boldsymbol{\sigma}) U(\gamma_1) = \\ &= \gamma_2\gamma_1(\boldsymbol{\sigma}) = U^+(\gamma_2\gamma_1) \boldsymbol{\sigma} U(\gamma_2\gamma_1) .\end{aligned}$$

Thus we have that the unitary operator $V = U(\gamma_2\gamma_1)U^+(\gamma_1)U^+(\gamma_2)$ is such that

$$V^+ \boldsymbol{\sigma} V = \boldsymbol{\sigma} \quad (10.6)$$

or

$$\boldsymbol{\sigma} V = V \boldsymbol{\sigma} . \quad (10.7)$$

Being V an element of $SU(2)$ we easily verify that it implies $V = 1$ or $V = -1$. We can then say that (10.2) gives a projective representation of $SO(3)$ i.e. (10.4) holds with the phase factor α which can take only the values ± 1 . This tells us also that if we have a sequence of transformations of $SO(3)$ whose product equals the identity, under the related $SU(2)$ transformations the spinor is transformed either in itself or just changes sign.

Vice-versa given an element U of $SU(2)$ we can write

$$U^+ \sigma_j U = \sum_{l=1}^3 \Gamma_{jl} \sigma_l \quad (10.8)$$

as the trace of the first member is zero. Being the first member an hermitean operator we have that the elements Γ_{jl} are real. Taking the product of two such relations and taking the trace we have

$$\delta_{jm} = \sum_{l=1}^3 \Gamma_{jl} \Gamma_{ml}$$

from which we conclude that the Γ_{jl} are elements of the real orthogonal group. Moreover if we take the trace of $U^+\sigma_1UU^+\sigma_2UU^+\sigma_3U$ we find

$$2i = 2i \sum_{jkl} \epsilon_{jkl} \Gamma_{1j} \Gamma_{2k} \Gamma_{3l}$$

i.e. $\det \Gamma = 1$. U and $-U$ through (10.8) generate the same Γ .

If U_1 and U_2 generate the same Γ due to the reasoning of eq.(10.6,10.7)) we have $U_2 = \pm U_1$. Thus to each element of $SU(2)$ there corresponds an element of $SO(3)$ while to each element of $SO(3)$ there correspond two elements $SU(2)$ given by $\pm U$.

$SU(2)$ is a simply connected group which is the universal covering group of $SO(3)$. This is seen by the fact that from (10.3) it is topologically equivalent to the surface of a four dimensional sphere which is also topologically equivalent to two tridimensional balls of radius 1 having in common the external surface.

Instead $SO(3)$ is topologically equivalent to a ball of radius π with the diametrically opposite points of the surface of such a ball, identified.

This is due to the fact that each element of $SO(3)$ is the rotation around an axis and the rotation around an axis by π is the same transformation which we obtain rotating by π around the opposite direction. As a result $SO(3)$ is not simply connected.

Chapter 11

Identical particles

11.1 Introduction

There are no observable differences when we exchange two identical particles. In quantum mechanics however some phenomena arise which have no counterpart in classical physics. Is it possible to distinguish two states which differ by the exchange of two particles?

Classically yes because the two particles can be distinguished by the different initial conditions. This presupposes the possibility to follow the particles along their trajectories.

This we know is not possible in quantum mechanics and it gives rise to the operative impossibility to distinguish two states in which e.g. two electrons have been exchanged. In a surprising and also satisfactory manner such the indistinguishable nature is more than operative but results in some new axioms in the theory which restrict the possible states when we are in presence of identical particles.

Let us consider n identical particles: we can describe the first particle with some dynamical variables ξ_1 (e.g. $\mathbf{q}_1, \mathbf{p}_1, \mathbf{s}_1$) the second with ξ_2 etc... . As they refer to different systems we shall have that the ξ_1 commutes with the ξ_2 etc... . The fact that the particles are identical is translated into the fact that the Hamiltonian will be a symmetric function of the variables ξ_1, ξ_2 etc.. i.e. invariant under the exchange of any pair of particle and this must happen whatever physical perturbation the system is subject to. All this is obviously true also in classical

physics. If we denote with $\psi_a^{(1)}, \psi_b^{(1)} \dots$ a basis of vectors which describe the first particle we shall have analogous vectors $\psi_a^{(2)}, \psi_b^{(2)} \dots$ which describe the second particle etc.... The Hilbert space $\mathcal{H}_n$ which describes the system of n particles is in natural manner given by the tensor product of the Hilbert spaces of the single particles; a basis of such tensor product is given by $\psi_a^{(1)} \psi_b^{(2)} \psi_c^{(3)} \dots$ and such a basis is called a symmetric basis. If we exchange two particles e.g. 1 with 2 the previously described state changes into $\psi_a^{(2)} \psi_b^{(1)} \psi_c^{(3)} \dots$. The mean values of physical observables, which are also symmetric under exchange of the dynamical variables $\xi_1, \dots, \xi_n$, do not change. Clearly factorized vectors of the above type do not exhaust the Hilbert space, in the sense that such a space will be formed by arbitrary superposition of such factorized vectors. The exchange of two particles generates a linear transformation in Hilbert space and it is immediate to realize that such operator which represents the exchange does not depend on the particular symmetric basis chosen in $\mathcal{H}_n$.

Thus given any vector $\psi \in \mathcal{H}_n$ we can consider the action of whatever permutation on it, $P\psi$ and we shall have a linear representation of the permutation group. Among the vectors of $\mathcal{H}_n$ there are some which under the action of P changes only by a phase, i.e. ψ and $P\psi$ represent the same state whatever $P \in S_n$ being S_n the group of permutation of n objects. In this case given such a ψ , the one dimensional space of vectors $\{P\psi\}$ is the basis of a one dimensional representation of S_n . It is a simple exercise to prove that there are only two possibilities: a) $P\psi = \psi$ for all $P \in S_n$; b) $P\psi = \delta_P \psi$ being $\delta_P = \pm 1$ according to the fact that P be an even or odd permutation and we have $\delta_{PQ} = \delta_P \delta_Q$. In the first case the vector ψ will be called symmetric and in the second will be called antisymmetric. The important fact is that such symmetric or antisymmetric property are conserved in time due to the symmetry with respect to permutations of the evolution operator, i.e. of the Hamiltonian H .

In fact as $P H = H P$ whatever P , we have $P H \psi = H P \psi = \pm H \psi$ and thus if $P \psi = \pm \psi$ we have $P(\psi + \frac{\epsilon}{i\hbar} H \psi) = \pm(\psi + \frac{\epsilon}{i\hbar} H \psi)$. Thus a state which is initially symmetric, stays symmetric, while one which is antisymmetric stays antisymmetric.

Thus there is the consistent possibility that for certain types of particles only

symmetric states are realized in nature, and for others only antisymmetric states are realized. This by the way would realize a very elegant formal peculiarity: indistinguishable states are described by the same ray in Hilbert space; in fact the exchange operator P changes at most the sign of the vector.

This is what happens in nature. One finds in fact that all experimental data agree with the following axiom: Particle with half-integer spin, called fermions, are described by vectors antisymmetric under exchange and those of integers spin called bosons, by vectors symmetric under the exchange of particles.

Let us see now how one can build in the Hilbert space $\mathcal{H}_n$ symmetric or antisymmetric states. If we start from the vectors $\psi_a^{(1)}, \psi_b^{(2)}, \dots, \psi_g^{(n)}$, with $a, b, c \dots$ not necessarily different among themselves, a symmetric vector is easily constructed with the operation

$$\psi_S = \sum_P P \psi_a^{(1)} \psi_b^{(2)} \dots \psi_g^{(n)} .$$

For a such a state we can only say that there are n_a particles in the state a , n_b particles in the state b etc.. without being able to specify which particle occupy such states. An antisymmetric vector is easily built in analogous manner

$$\psi_A = \sum_P \delta_P P \psi_a^{(1)} \psi_b^{(2)} \dots \psi_g^{(n)}$$

which is equivalent to the determinant

$$\begin{vmatrix} \psi_a^{(1)} & \psi_a^{(2)} & \dots & \psi_a^{(n)} \\ \psi_b^{(1)} & \psi_b^{(2)} & \dots & \psi_b^{(n)} \\ \dots & \dots & \dots & \dots \\ \psi_g^{(1)} & \psi_g^{(2)} & \dots & \psi_g^{(n)} \end{vmatrix} .$$

An immediate result is that if two indices e.g. a and b are equal, due to the antisymmetry we have $\psi_A = 0$. This can be also seen from the fact that we must have at the same time

$$P_{12}\psi_A = \psi_A$$

and

$$P_{12}\psi_A = -\psi_A .$$

Thus there cannot be two identical fermions in the same state. The states described above are not the only symmetric or antisymmetric states in $\mathcal{H}_n$.

Linear combinations of symmetric or antisymmetric states of the type above described cannot be written in the above form even if they maintain the property of symmetry or antisymmetry. The physical Hilbert space is given by $\mathcal{H}_n^S$ or by $\mathcal{H}_n^A$ whether the particles are boson or fermions.

This axiom has very important consequences on the statistics of particle systems, i.e. on the counting of states and thus on the computations of the mean values which occur in statistical thermodynamics.

In fact according to the particles being bosons or fermions, the traces which give rise to the statistical mean values of observables, are to be taken on the symmetric subspace of $\mathcal{H}$ i.e. $\mathcal{H}_S$, or on the antisymmetric subspace $\mathcal{H}_A$.

$$\langle F \rangle = \text{Tr}_{\mathcal{H}_S}(F e^{-\beta H}) / \text{Tr}_{\mathcal{H}_S}(e^{-\beta H}); \quad \langle F \rangle = \text{Tr}_{\mathcal{H}_A}(F e^{-\beta H}) / \text{Tr}_{\mathcal{H}_A}(e^{-\beta H}) .$$

Let us notice that in completely general way, i.e. independently of the principle of spin and statistics, a permutation $P \in S_n$ induces a unitary transformation in the Hilbert space $\mathcal{H}_n$; this is seen from how P operates on the vectors of a symmetric basis for which

$$(P\psi_{a'}^{(1)}\psi_{b'}^{(2)}\dots\psi_{g'}^{(n)}, P\psi_a^{(1)}\psi_b^{(2)}\dots\psi_g^{(n)}) = (\psi_{a'}^{(1)}\psi_{b'}^{(2)}\dots\psi_{g'}^{(n)}, \psi_a^{(1)}\psi_b^{(2)}\dots\psi_g^{(n)}) .$$

It is known that the interaction of the spin of a particle with the other degrees of freedom and with the spin of other particles is a relativistic phenomenon and that in light atoms such interaction is estimated of the order $(\frac{1}{137})^2$. Thus it is a very good procedure to neglect in zero order approximation, the spin in the Hamiltonian and to introduce the interactions due to the spin later as a perturbation.

Given now a Hamiltonian which depends only on the coordinates and conjugate momenta, we want to examine the symmetry properties under permutations of the energy eigensolutions of the Schrödinger equation, independently of the principle of spin and statistics. The results will be combined later with the spin degrees of freedom in such a way as to obtain symmetric or antisymmetric state vectors.

11.2 Two particle system

Let us consider in the case of two particles a solution of the Schrödinger equation

$$H\psi(q_1, q_2) = E\psi(q_1, q_2) .$$

The identity of the two particle implies that the Hamiltonian commutes with the elements of the permutation group which in this simple case are two i.e. the identity and P_{12} . Such a symmetry allows us to derive from $\psi(q_1, q_2)$ an other solution. In fact

$$P_{12}H\psi(q_1, q_2) = HP_{12}\psi(q_1, q_2) = EP_{12}\psi(q_1, q_2) .$$

Let us notice that such a vector $P_{12}\psi(q_1, q_2)$ cannot be the null vector as $P_{12}^2 = 1$. If the new vector $P_{12}\psi(q_1, q_2)$ is not proportional to $\psi(q_1, q_2)$ we have at our disposal a two dimensional subspace of solutions to the energy E . It is very useful to see such a two dimensional space as the direct sum of two one dimensional spaces invariant under the group P . In fact after forming the two vectors

$$\psi(q_1, q_2) + P_{12}\psi(q_1, q_2), \quad \psi(q_1, q_2) - P_{12}\psi(q_1, q_2)$$

we have that

$$P_{12}\left(\psi(q_1, q_2) \pm P_{12}\psi(q_1, q_2)\right) = \pm\left(\psi(q_1, q_2) \pm P_{12}\psi(q_1, q_2)\right)$$

and thus the two one dimensional spaces generated by $\psi(q_1, q_2) \pm P_{12}\psi(q_1, q_2)$ are invariant.

Obviously it may happen that the starting $\psi(q_1, q_2)$ be already symmetric or antisymmetric with respect to the exchange in which case one of the two vectors $\psi(q_1, q_2) \pm P_{12}\psi(q_1, q_2)$ vanishes and we have only one one-dimensional space. In the general case of n identical particles given a solution of the Schrödinger equations

$$H\psi(q_1, q_2, \dots q_n) = E\psi(q_1, q_2, \dots q_n)$$

we can construct a linear space of solutions by applying to this solution the permutation group

$$P\psi(q_1, q_2, \dots q_n) .$$

In this way we obtain a space which is at most $n!$ dimensional and again it is important to decompose such a linear space into subspaces invariant under the permutation group.

In the two particle case we combine the obtained wave functions with the spin vectors in order to satisfy the axiom of spin and statistics. Using the Young *diagrams* we denote by

$$\begin{array}{|c|c|} \hline & \\ \hline \end{array} \psi(q_1, q_2) \quad (11.1)$$

the symmetrized wave function and by

$$\begin{array}{|c|} \hline \\ \hline \\ \hline \end{array} \psi(q_1, q_2) \quad (11.2)$$

the antisymmetrized one. Then in the case of two identical fermions of spin $1/2$ as the vectors of total spin 1 i.e

$$|++\rangle, \quad \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle), \quad |--\rangle$$

are symmetric under exchange these can be multiplied only by the antisymmetric wave function (11.2), while being the spin state of total spin 0 antisymmetric

$$\frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle)$$

it can be multiplied only by the symmetric wave function (11.1).

11.3 Three particle system

Let us consider the simple but not trivial case of three particles. An arbitrary wave function $\psi(q_1, q_2, q_3)$ can be rewritten as the sum of the following four wave functions

$$\begin{aligned} \psi(q_1, q_2, q_3) = & \begin{array}{|c|c|c|} \hline & & \\ \hline \end{array} \psi(q_1, q_2, q_3) + \begin{array}{|c|} \hline \\ \hline \\ \hline \end{array} \psi(q_1, q_2, q_3) + \\ & \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \end{array} \psi(q_1, q_2, q_3) + \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & \end{array} \psi(q_1, q_2, q_3) = \end{aligned}$$

$$\begin{aligned}
& \frac{1}{6}[(123) + (132) + (213) + (231) + (312) + (321)] \\
& + \frac{1}{6}[(123) - (132) - (213) + (231) + (312) - (321)] \\
& + \frac{1}{3}[(123) + (213) - (321) - (231)] \\
& + \frac{1}{3}[(123) + (321) - (213) - (312)]
\end{aligned}$$

The Young *diagram*

--	--	--

 is the symmetrizer, the Young *diagram*

 is the antisymmetriser

while the Young *table*

1	2
3	

 means, first symmetrize with respect to 1 and 2

and then antisymmetrize the found result with respect to 1 and 3.

In analogous manner the *table*

1	3
2	

 means, first symmetrize with respect to

1 and 3 and then antisymmetrize the found result with respect to 1 and 2.

If we apply an element of the permutation group to

--	--	--

 ψ we find the same vector and thus $\{P\}$

--	--	--

 ψ is a one dimensional space. Similarly $\{P\}$

 ψ gives rise to a one dimensional representation

$$Q \sum_P \delta_P P \psi = \sum_P \delta_P Q P \psi = \delta_Q \sum_P \delta_{PQ} Q P \psi = \delta_Q \sum_P \delta_P P \psi.$$

I.e. if we apply an element of the permutation group we find the same vector multiplied by 1 or -1 for even or odd permutation.

Let us examine now the subspace $\{P\}\alpha$ with $\alpha = (123) + (213) - (321) - (231)$.

We have

$$\begin{aligned} P_{12}\alpha &= P_{12}[(123) + (213) - (321) - (231)] = \\ &= (213) + (123) - (312) - (132) \equiv \beta \end{aligned}$$

$$\begin{aligned} P_{12}\beta &= P_{12}[(213) + (123) - (312) - (132)] = \\ &= (123) + (213) - (321) - (231) = \alpha \end{aligned}$$

as $P_{12}^2 = 1$.

$$\begin{aligned} P_{13}\alpha &= P_{13}[(123) + (213) - (321) - (231)] = \\ &= (321) + (231) - (123) - (213)] = -\alpha \end{aligned}$$

$$\begin{aligned} P_{13}\beta &= P_{13}[(213) + (123) - (312) - (132)] = \\ &= (231) + (321) - (132) - (312)] = -\alpha + \beta . \end{aligned}$$

P_{12} and P_{13} generate the whole permutation group of three elements: in fact

$$P_{12}(123) = (213); \quad P_{13}(213) = (231); \quad P_{12}(231) = (132);$$

$$P_{13}(132) = (312); \quad P_{12}(312) = (321) .$$

In the basis $v_1 = \alpha$, $v_2 = \beta$ we have the representation

$$Pv_i = \sum_j v_j P(j, i)$$

with

$$P_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P_{13} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} .$$

From these by multiplication we can compute the representatives of all permutations.

In similar way we can compute the representation generated by $\begin{bmatrix} 1 & 3 \\ 2 \end{bmatrix} \psi \equiv \gamma$.

$$P_{12}\gamma = -\gamma, \quad P_{13}\gamma \equiv \epsilon, \quad P_{12}\epsilon = -\gamma + \epsilon, \quad P_{13}\epsilon = \gamma$$

and we find the representation

$$P'_{12} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \quad P'_{13} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

The two representations are equivalent by a similitude transformation. In fact given

$$S = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}$$

we have $PS = SP'$ which is enough to verify for P_{12} and P_{13} . This is a general fact: representations generated by Young *tables* having the same Young *diagram* are equivalent. It is a simple exercise to show that the space spanned by α and β is either two-dimensional or it is the null space, i.e. if we set $\beta = c\alpha$ we have $\alpha = 0$ and that these two representations (i.e. α, β and γ, ϵ), are irreducible, i.e. if the subspace $\alpha + c\beta$ were invariant we have $\alpha = \beta = 0$. This is a general result: spaces generated applying the permutation group to a vector symmetrized by means of a Young symmetrizer are irreducible. The two subspaces (α, β) and (γ, ϵ) are not always independent, in the sense that one can be the null subspace, or it may happen that the two subspaces coincide. This happens e.g. if the starting wave function satisfies $\psi(q_1, q_2, q_3) = \psi(q_1, q_3, q_2)$.

One verifies however that vectors belonging to spaces generated by different Young *diagrams* (i.e. symmetric, antisymmetric and mixed) are orthogonal. This is also a general result.

The fact that wave functions belonging to irreducible inequivalent representations, i.e. to different Young *diagrams* are orthogonal is a general consequence of the orthogonality of inequivalent representations

$$(\psi_i, \phi_j) = \frac{\sum_P (P\psi_i, P\phi_j)}{n!} = \frac{1}{n!} \sum_P \sum_k \sum_h P^*(k, i) P'(h, j) (\psi_k, \phi_h)$$

where $P(k, i)$ is the representation to which ψ_i belong and $P'(k, i)$ is the representation to which ϕ_j belong. But the theorem on the orthogonality of the matrix elements of irreducible inequivalent representations tells us

$$\sum_P P^*(k, i) P'(h, j) = 0$$

and thus $(\psi_i, \phi_j) = 0$.

We saw that given a generic wave function, by applying to it the permutation group we obtain an $n!$ dimensional space, in our example a 6 dimensional space. Notice that we have three inequivalent representations: two inequivalent one dimensional representations and two equivalent two dimensional representations. Even this is a general result: In the decomposition of the generic vector in irreducible representations each representation appears as many times as the dimension of the representation itself and thus (Burnside theorem) the sum of the squares of the dimensions of inequivalent representations equals the dimension of the group i.e. $n!$ in the case of S_n .

In the case in which the wave function is not generic but e.g. a stationary solution of a specific Schrödinger equation, is not granted that the space generated by applying the permutation group be $n!$ dimensional, in our case 6 dimensional.

We want now to examine which are the possible dimensions of such a space for a generic Hamiltonian invariant under the permutation group. I.e. let us suppose that initially the space of an energy level be effectively 6 times degenerate.

We ask in how many levels and with which degeneracy such a level splits under the action of an arbitrary perturbation whose only requirements will be to be hermitean and invariant under permutations. If a perturbation existed non invariant under permutations we would have discovered a method to distinguish the particles which would be no longer identical.

We have the following simple result: If $\{P\}$ is a group of unitary transformations in $\mathcal{H}$ and V is a subspace invariant under such group of transformations, also its orthogonal complement $V_\perp$ is invariant under such a group of transformations. In fact if $w \in V_\perp$ for any $v \in V$ and any P , we have $(v, Pw) = (P^+v, w) = (P^{-1}v, w) = 0$ because $P^{-1}v \in V$ and thus $Pw \in V_\perp$.

In this way we can decompose $\mathcal{H}$ successively in orthogonal invariant subspaces to reach at the end a decomposition in invariant orthogonal irreducible subspaces.

Moreover if V is invariant the projector on V , $\mathcal{P}_V$ commutes with all the P . In fact chosen an orthogonal basis v_i in such a subspace the projector can be written

as $\mathcal{P}_V = \sum_i v_i \circ v_i$ and we have

$$P\mathcal{P}_VP^{-1} = P\mathcal{P}_VP^+ = \sum_i P v_i \circ P v_i = \sum_{ikl} P(k, i) P^*(l, i) v_k \circ v_l = \sum_i v_i \circ v_i$$

due to the unitary nature of the transformation P .

The simplest Hamiltonian which breaks the degeneracy is given by a real combination of projectors on the invariant orthogonal subspaces. We saw how the symmetric vector generates an invariant subspace which is orthogonal to the one generated by the antisymmetric vector which at the same time are orthogonal to the four dimensional space generated by the two mixed Young *tables*.

Let us consider within the four-dimensional space generated by α and γ any invariant two dimensional subspace. We know that in the mentioned four dimensional space there is no invariant subspace of dimension one.

The orthogonal complement, within the four dimensional space, of such two dimensional space is also an invariant subspace.

We already saw that if we have a subspace V invariant under the permutation group the projector on such subspace commutes with all permutations.

In this way the four projectors on the described invariant orthogonal subspaces constitute four hermitean operators, which commute with all permutations and thus an acceptable perturbation Hamiltonian is

$$c_1\mathcal{P}_1 + c_2\mathcal{P}_2 + c_3\mathcal{P}_3 + c_4\mathcal{P}_4$$

with real arbitrary c_i . Such a perturbation breaks the degenerate level into four levels.

Thus a generic perturbation Hamiltonian, invariant under the permutation group breaks the original 6 times degenerate level in four distinct levels, of which two non degenerate, i.e. the symmetric and the antisymmetric ones, and two doubly degenerate.

Such a residual degeneracy cannot be further removed. In fact within the two dimensional spaces according to perturbation theory we have to compute the eigenvalues of $(v_i, H_I v_j)$ where H_I is the perturbation. But we must have $PH_I = H_IP$ and we have

$$(v_i, H_IP v_j) = \sum_k (v_i, H_I v_k) P(k, j)$$

and

$$(v_i, PH_I v_j) = (P^+ v_i, H_I v_j) = \left(\sum_l P^*(i, l) v_l, H_I v_j \right) = \sum_l P(i, l) (v_l, H_I v_j) .$$

Thus being $P(i, l)$ irreducible due to Schur lemma we have $(v_i, H_I v_j) = c \delta_{ij}$ and it is not possible to remove the degeneracy.

Notice that the above analysis is completely independent of the axiom of spin and statistics.

It is now important to study the structure of the state vector in presence of such axiom.

I.e. we must find which vectors in the space of three identical particles are totally symmetric or totally antisymmetric.

If the particles are of spin zero and as such bosons of the studied wave functions the only acceptable one is $\begin{bmatrix} \square & \square & \square \end{bmatrix} \psi$.

We consider now the case of three identical particles of spin 1/2 (fermions).

A spinor of total spin 3/2 is symmetric under the exchange of particles and thus the only way to form a totally antisymmetric vector is to multiply it by the unique antisymmetric function we have at our disposal i.e. $\begin{bmatrix} \square \\ \square \\ \square \end{bmatrix} \psi$.

Suppose now to have a combination of the vectors α and β with spinors such that the vector is totally antisymmetric.

Without loosing generality we can consider states with well defined S_z . In fact we can project our state on the eigensubspaces of S_z using the projectors on such subspaces e.g. the projector $-(S_z - 3/2)(S_z + 1/2)(S_z + 3/2)/2$, which being symmetric in the spin operators does not alter the symmetry nature of the vector. We know already that being $S_z = \pm 3/2$ completely symmetric it cannot be used. Thus we use $S_z = \pm 1/2$, referring for concreteness to the case $S_z = 1/2$. The most general vector with $S_z = 1/2$ is

$$\psi = [a | ++- \rangle + b | +-+ \rangle + c | -++ \rangle] \alpha + [d | +-+ \rangle + e | +-- \rangle + f | -++ \rangle] \beta.$$

Being P_{12} odd, we must have

$$P_{12}\psi = [a |++-\rangle + b |-++\rangle + c |+-+\rangle]\beta + [d |++-\rangle + e |-++\rangle + f |+-+\rangle]\alpha = -\psi$$

and thus comparing the two expressions we have $-a = d$, $-c = e$, $-b = f$, $-d = a$, $-e = c$, $-f = b$. It follows

$$\psi = [a |++-\rangle + b |+-+\rangle + c |-++\rangle]\alpha - [a |++-\rangle + c |+-+\rangle + b |-++\rangle]\beta.$$

Moreover

$$P_{13}\psi = -[a |-++\rangle + b |+-+\rangle + c |++-\rangle]\alpha - [a |-++\rangle + c |+-+\rangle + b |++-\rangle](-\alpha + \beta) = -\psi$$

from which $-a = -c + b$, $-c = -a + a = 0$, $a = -b$, which gives us at last

$$\psi = [|++-\rangle - |-++\rangle]\alpha - [|++-\rangle - |-++\rangle]\beta.$$

As from P_{12} and P_{13} we can compose all permutations the total antisymmetry of ψ is proven.

Notice that with $S_+ = S_1 + iS_2$ we have

$$S_+ (|-++\rangle - |-++\rangle) = 0$$

as well

$$S_+ (|++-\rangle - |-++\rangle) = 0$$

i.e. both spin vectors are eigenstates of total spin $S = 1/2$.

Thus we found that for two or three electrons (particles of spin $1/2$) to each value of the total spin there corresponds a well defined irreducible representation of the permutation group for the space wave functions.

This is also a general result for n particles of spin $1/2$ which by the spin-statistics theorem obey Fermi statistics.

References

- [1] E.P. Wigner *Group theory*, Academic Press, New York, London (1959)
- [2] H. Weyl *The theory of groups and quantum mechanics*, Dover Publications, New York (1950)

Chapter 12

The formal theory of scattering

12.1 Introduction

The elementary treatment of potential scattering is performed in the time independent method: one looks for a stationary solution of the Schrödinger equation which obeys proper boundary conditions. The computation of the current at large distances allows to go over from the so computed wave function to the cross section. There exists a variety of more complex scattering problems for which such a formulation is not apt; the time dependent method in which one explicitly considers the evolution of the state appears more useful, both in deriving general properties of the scattering states and of transitions amplitudes and to perform approximate computation, like perturbative expansions. In such time dependent methods one tries to reproduce in a mathematical language a situation more or less similar to the experimental setting, which is characterized by the fact the before and after the scattering process, the intervening particles (subsystems) are free. This can be obtained in two ways 1) Work with wave packets (obviously of not perfectly defined energy) and then analyze the results in terms of energy eigenstates. 2) Use states of perfectly defined energy (plane wave) as initial states and switch off artificially the interaction before and after a certain time interval and then have such time interval tend to infinity as to have a process of well defined energy. Clearly the second method is more artificial than the one dealing with wave packets, but formally simpler. After the switching off of the interac-

tion the particle is again free and the momentum is a constant of motion; one analyzes such state in momentum eigenstates obtaining in this way the transition amplitude.

In the present treatment we use the second method, following the lines of the classical paper by Lippmann and Schwinger [1]; the treatment by means of wave packets has been given by Gell-Mann and Goldberger [2] and in mathematical rigorous way in Jauch [3]

12.2 The Lippmann-Schwinger equation

Let us consider the evolution in time of a system composed of two interacting parts. Let H be the hamiltonian of the system $H = H_0 + V$ where H_0 is the free hamiltonian and V the interaction.

The generic state vector $|\psi^s(t)\rangle$ satisfies Schrödinger's equation

$$i\frac{\partial}{\partial t}|\psi^s(t)\rangle = (H_0 + V)|\psi^s(t)\rangle$$

where t here and in the following stays for $t/\hbar$. As we want to describe the effect of V , to remove the time dependence associated to H_0 let us perform the transformation

$$|\psi(t)\rangle = e^{iH_0 t}|\psi^s(t)\rangle$$

thus obtaining

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = H_1(t)|\psi(t)\rangle, \quad H_1(t) = e^{iH_0 t}V e^{-iH_0 t}.$$

The evolution operator defined by

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$$

with the property $U(t, t) = 1$ satisfies

$$i\frac{\partial}{\partial t}U(t, t_0) = H_1(t)U(t, t_0). \quad (12.1)$$

Such a differential equation can be rewritten in the form of an integral equation

$$U(t_2, t_1) = 1 - i \int_{t_1}^{t_2} H_1(t')U(t', t_1)dt'$$

which includes the condition $U(t_1, t_1) = 1$. The main properties of $U(t_2, t_1)$ have been discussed in Chapter 4.

Eq.(12.1) can be rewritten as

$$U(t, -\infty) = 1 - i \int_{-\infty}^t H_1(t') U(t', -\infty) dt' \quad (12.2)$$

provided the limit for $t_1 \rightarrow -\infty$ exists. We are interested in the transition between eigenstates of the free hamiltonian $|\phi_a\rangle, |\phi_b\rangle$ with

$$H_0|\phi_a\rangle = E_a|\phi_a\rangle, \quad H_0|\phi_b\rangle = E_b|\phi_b\rangle \quad a \neq b.$$

We shall introduce following the general ideas discussed at the beginning of this chapter, an adiabatic switch-off factor, $V \rightarrow e^{-\varepsilon|t|}V$ and thus we have the transition amplitude

$$T_{ba} = \langle \phi_b | -i \int_{-\infty}^{+\infty} e^{iH_0t} e^{-\varepsilon|t|} V e^{-iH_0t} U(t, -\infty) dt | \phi_a \rangle.$$

A slow switching on and off of the interaction is necessary because only with an adiabatic process the energy of the final state will equal that of the initial state, something which is not true for a sudden switch-on or switch-off. We shall show this in the following. We have

$$T_{ba} = -i \langle \phi_b | V | \Psi_a^+(E_b) \rangle$$

with

$$|\Psi_a^+(E_b)\rangle = \int_{-\infty}^{\infty} e^{-\varepsilon|t|} e^{iE_b t} e^{-iH_0 t} U(t, -\infty) |\phi_a\rangle dt.$$

We want to obtain an operator equation for such a state which will be at the center of our interest. From eq.(12.2) we obtain

$$\begin{aligned} |\Psi_a^+(E_b)\rangle &= \int_{-\infty}^{+\infty} e^{-\varepsilon|t|} e^{iE_b t} e^{-iH_0 t} |\phi_a\rangle dt \\ &\quad - i \int_{-\infty}^{+\infty} e^{-\varepsilon|t|} e^{iE_b t} e^{-iH_0 t} dt \int_{-\infty}^t H_I(t') U(t', -\infty) |\phi_a\rangle dt' \\ &= 2\pi\delta_\varepsilon(E_b - E_a) |\phi_a\rangle \\ &\quad - i \int_0^\infty d\tau e^{iE_b \tau} e^{-iH_0 \tau} e^{-\varepsilon\tau} V \int_{-\infty}^{+\infty} e^{iE_b t'} e^{-iH_0 t'} [e^{\varepsilon\tau} e^{-\varepsilon|\tau+t'|} e^{-\varepsilon|t'|}] U(t', -\infty) |\phi_a\rangle dt' \end{aligned}$$

obtained by setting $\tau = t - t'$ and where

$$\delta_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2} .$$

The switching-off factor in the square bracket differs from $e^{-\varepsilon|t'|}$ for the fact that it stays constant equal to 1 in the interval $-\tau < t' < 0$ and at infinity decays as $e^{-2\varepsilon|t'|}$. It is an assumption of the treatment that for $\varepsilon \rightarrow 0$ the two adiabatic switching off factors are equivalent. We have then

$$|\Psi_a^+(E_b)\rangle = 2\pi\delta_\varepsilon(E_b - E_a)|\phi_a\rangle + \frac{1}{E_b - H_0 + i\varepsilon}V|\Psi_a^+(E_b)\rangle .$$

To understand the nature of this equation let us take the scalar product with a complete set of vectors $|n\rangle$.

$$\langle n|\Psi_a^+(E_b)\rangle = 2\pi\delta_\varepsilon(E_b - E_a)\langle n|\phi_a\rangle + \frac{1}{E_b - E_n + i\varepsilon} \int \langle n|V|m\rangle\langle m|\Psi_a^+(E_b)\rangle dm .$$

This is an equation of type

$$F(x) = f(E_a, E_b)G(x) + \frac{1}{E_b - E(x) + i\varepsilon} \int V(x, x')F(x')dx'$$

where $f(E_a, E_b)$ is a multiplicative factor inessential to the solution. Therefore it is meaningful to divide by $\delta_\varepsilon(E_b - E_a)$, i.e. consider a new vector $|\psi_a^+\rangle$ given by

$$|\Psi_a^+(E_b)\rangle = 2\pi\delta_\varepsilon(E_b - E_a)|\psi_a^+\rangle \quad (12.3)$$

to obtain

$$|\psi_a^+\rangle = |\phi_a\rangle + \frac{1}{E_a - H_0 + i\varepsilon}V|\psi_a^+\rangle . \quad (12.4)$$

This is the Lippmann-Schwinger equation. The $\delta_\varepsilon(E_b - E_a)$ in eq.(12.3) clearly indicates energy conservation in the limit $\varepsilon \rightarrow 0$. The obtained $|\psi_a^+\rangle$ is a solution of Schrödinger equation at the energy E_a . In fact

$$(E_a - H_0)|\psi_a^+\rangle = (E_a - H_0)|\phi_a\rangle + V|\psi_a^+\rangle$$

i.e.

$$(E_a - H)|\psi_a^+\rangle = 0 .$$

Thus starting from a time dependent equation we obtain in the limit $\varepsilon \rightarrow 0$ a time independent equation which defines a particular solution of the Schrödinger equation.

$|\psi_a^+\rangle$ satisfies also an other equation which will become useful in several formal developments. Multiplying $|\psi_a^+\rangle$ by V we obtain

$$\begin{aligned} V|\psi_a^+\rangle &= V|\phi_a\rangle + V \frac{1}{E_a - H_0 + i\varepsilon} V|\psi_a^+\rangle \\ (1 - V \frac{1}{E_a - H_0 + i\varepsilon}) V|\psi_a^+\rangle &= V|\phi_a\rangle \end{aligned}$$

or

$$(E_a - H + i\varepsilon) \frac{1}{E_a - H_0 + i\varepsilon} V|\psi_a^+\rangle = V|\phi_a\rangle$$

i.e.

$$|\psi_a^+\rangle = |\phi_a\rangle + \frac{1}{E_a - H + i\varepsilon} V|\phi_a\rangle .$$

We prove now the orthogonality relation for the $|\psi_a^+\rangle$

$$\begin{aligned} \langle \psi_b^+ | \psi_a^+ \rangle &= (\langle \phi_b | + \langle \phi_b | V \frac{1}{E_b - H - i\varepsilon}) | \psi_a^+ \rangle \quad (12.5) \\ &= \langle \phi_b | \psi_a^+ \rangle + \langle \phi_b | V \frac{1}{E_b - H - i\varepsilon} | \psi_a^+ \rangle = \langle \phi_b | \psi_a^+ \rangle + \frac{1}{E_b - E_a - i\varepsilon} \langle \phi_b | V | \psi_a^+ \rangle \\ &= \langle \phi_b | \phi_a \rangle + \langle \phi_b | \frac{1}{E_a - H_0 + i\varepsilon} V | \psi_a^+ \rangle + \frac{1}{E_b - E_a - i\varepsilon} \langle \phi_b | V | \psi_a^+ \rangle = \langle \phi_b | \phi_a \rangle . \end{aligned}$$

12.3 The adiabatic theorem

We prove now the “adiabatic theorem”: The solution $|\psi^+\rangle$ of Schrödinger’s equation, characterized by

$$|\psi^+\rangle = |\phi\rangle + \frac{1}{E - H_0 + i\varepsilon} V|\psi^+\rangle, \quad (H - E)|\psi^+\rangle = 0$$

when the coordinate representation makes sense, satisfy the usual boundary conditions which we impose in the elementary treatment of non relativistic scattering, i.e. at space infinity only outgoing waves are present. To this end we calculate

$$\begin{aligned} \langle \mathbf{r} | \frac{1}{E + i\varepsilon - H_0} V | \psi^+ \rangle &= \int \frac{\langle \mathbf{r} | \mathbf{k}' \rangle \langle \mathbf{k}' | V | \psi^+ \rangle}{E + i\varepsilon - \frac{\mathbf{k}'^2}{2m}} d\mathbf{k}' \\ &= \frac{1}{(2\pi)^{3/2}} \int_0^\infty \frac{k'^2 dk'}{E + i\varepsilon - \frac{k'^2}{2m}} \int_{-1}^{+1} e^{ik'r\zeta} f(\zeta, k') d\zeta \quad (12.6) \end{aligned}$$

with

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad \frac{\mathbf{k}^2}{2m} = \frac{k^2}{2m} = E$$

where m stays for the mass M divided by $\hbar^2$ and

$$f(\zeta, k') = \int_{\zeta=\text{const}} \langle \mathbf{k}' | V | \psi^+ \rangle d\varphi$$

having assumed the direction of $\mathbf{r}$ as polar axis.

We have for the l.h.s. of eq.(12.6)

$$\frac{1}{(2\pi)^{3/2}} \int_0^\infty \left[\frac{e^{ik'r}}{ik'r} f(1, k') - \frac{e^{-ik'r}}{ik'r} f(-1, k') - \int_{-1}^1 \frac{e^{ik'r\zeta}}{ik'r} \frac{\partial f(\zeta, k')}{\partial \zeta} d\zeta \right] \frac{k'^2 dk'}{E + i\varepsilon - \frac{k'^2}{2m}}.$$

By the Riemann-Lebesgue theorem the term $\int_{-1}^1$ goes to zero more rapidly than $1/r$. Thus it is of no interest in the discussion of the asymptotic wave. We are left with

$$\begin{aligned} & \frac{1}{(2\pi)^{3/2}ir} \int_0^\infty \frac{e^{ik'r} f(1, k') - e^{-ik'r} f(-1, k')}{E + i\varepsilon - \frac{k'^2}{2m}} k' dk' = \\ & \frac{1}{(2\pi)^{3/2}ir} \left[f(1, k) \int_0^\infty \frac{e^{ik'r} k' dk'}{E + i\varepsilon - \frac{k'^2}{2m}} + \int_0^\infty \frac{e^{ik'r} (f(1, k') - f(1, k)) k' dk'}{E + i\varepsilon - \frac{k'^2}{2m}} \right. \\ & \left. - f(-1, k) \int_0^\infty \frac{e^{-ik'r} k' dk'}{E + i\varepsilon - \frac{k'^2}{2m}} - \int_0^\infty \frac{e^{-ik'r} (f(-1, k') - f(-1, k)) k' dk'}{E + i\varepsilon - \frac{k'^2}{2m}} \right]. \end{aligned}$$

Again by Riemann-Lebesgue theorem the only terms which survive at large r are

$$\frac{1}{(2\pi)^{3/2}ir} \left[f(1, k) \int_0^\infty \frac{e^{ik'r} k' dk'}{E + i\varepsilon - \frac{k'^2}{2m}} - f(-1, k) \int_0^\infty \frac{e^{-ik'r} k' dk'}{E + i\varepsilon - \frac{k'^2}{2m}} \right].$$

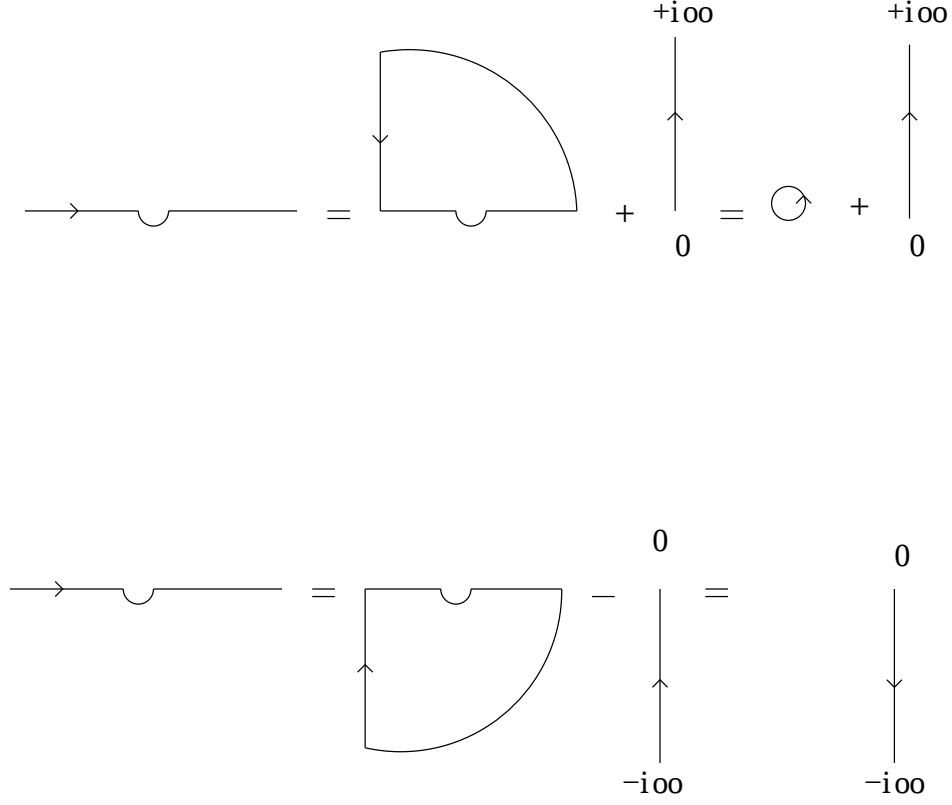
These terms can be computed by a contour integral in the complex plane.

$$\frac{1}{(2\pi)^{3/2}ir} \left\{ f(1, k) \left[-me^{ikr} 2\pi i - \int_0^\infty \frac{e^{-\chi r} \chi d\chi}{E + \frac{\chi^2}{2m}} \right] + f(-1, k) \int_0^\infty \frac{e^{-\chi r} \chi d\chi}{E + \frac{\chi^2}{2m}} \right\}$$

where we applied the Jordan lemma [4] to the contours depicted in figure 12.1.

The terms

$$\int_0^\infty \frac{e^{-\chi r} \chi d\chi}{E + \frac{\chi^2}{2m}}$$

Figure 12.1: Contours in the k complex plane

vanish for large r and we are left with

$$\langle \mathbf{r} | \frac{1}{E - H_0 + i\varepsilon} V | \psi^+ \rangle \sim -\frac{e^{ikr}}{r} \frac{2\pi i m}{(2\pi)^{3/2} i} f(1, k) = -\frac{(2\pi)^2 m}{(2\pi)^{3/2}} \frac{e^{ikr}}{r} \langle \phi_{\mathbf{k}} | V | \psi^+ \rangle$$

where

$$\langle \mathbf{r} | \phi_{\mathbf{k}} \rangle = \frac{1}{\sqrt{(2\pi)^3}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad \mathbf{k} = \frac{\mathbf{r}}{r} \sqrt{2mE}.$$

Recalling the elementary treatment of scattering where the wave function for large r is chosen

$$e^{i\mathbf{k} \cdot \mathbf{r}} + g(\theta, \phi) \frac{e^{ikr}}{r}$$

we have for the cross section

$$\sigma(\theta, \phi) = |g(\theta, \phi)|^2 = (2\pi)^4 m^2 |\langle \phi_{\mathbf{k}} | V | \psi^+ \rangle|^2 = \left(\frac{2\pi}{\hbar} \right)^4 M^2 |\langle \phi_{\mathbf{k}} | V | \psi^+ \rangle|^2. \quad (12.7)$$

12.4 “In” and “out” states

We proceed with the interpretation of $|\psi_a^+\rangle$ establishing the fundamental relation

$$|\psi_a^+\rangle = U(0, -\infty)|\phi_a\rangle .$$

We obtain this result by noticing that

$$|\Psi_a^+(E_b)\rangle = \int_{-\infty}^{+\infty} e^{iE_b t} e^{-\varepsilon|t|} e^{-iH_0 t} U(t, -\infty) dt = 2\pi\delta_\varepsilon(E_b - E_a)|\psi_a^+\rangle$$

is the Fourier transform of $A = e^{-\varepsilon|t|} e^{-iH_0 t} U(t, -\infty)|\phi_a\rangle$ at the frequency E_b which gives a $2\pi\delta_\varepsilon(E_b - E_a)$ if and only if

$$e^{-iH_0 t} U(t, -\infty)|\phi_a\rangle = e^{-iE_a t} |\psi_a^+\rangle .$$

For $t = 0$ we have

$$U(0, -\infty)|\phi_a\rangle = |\psi_a^+\rangle . \quad (12.8)$$

It means that if we start from a plane wave $|\phi_a\rangle$ at $t = -\infty$ and we switch on the interaction up to time 0 we obtain the scattering wave. We can also perform the symmetric treatment from $+\infty$ to $-\infty$ defining

$$|\Psi_a^-(E_b)\rangle = \int_{-\infty}^{+\infty} e^{iE_b t} e^{-\varepsilon|t|} e^{-iH_0 t} U(t, +\infty)|\phi_a\rangle dt .$$

We shall call ψ^+ “in state” and ψ^- “out state”; ψ^- is obtained replacing $-\infty$ with $+\infty$ in the definition of the ψ^+ . The out states satisfy an equation which is similar to the Lippmann-Schwinger equation (12.4) replacing $i\varepsilon$ with $-i\varepsilon$ as can be verified by repeating the procedure.

Summing up we have

$$\begin{aligned} |\psi_a^\pm\rangle &= |\phi_a\rangle + \frac{1}{E_a - H_0 \pm i\varepsilon} V |\psi_a^\pm\rangle \\ |\psi_a^\pm\rangle &= |\phi_a\rangle + \frac{1}{E_a - H \pm i\varepsilon} V |\phi_a\rangle . \end{aligned}$$

The orthogonality relation (12.5) can be extended to the out states

$$\langle\psi_b^\pm|\psi_a^\pm\rangle = \langle\phi_b|\phi_a\rangle .$$

We compute now $\langle \phi_b | U(+\infty, -\infty) | \phi_a \rangle$

$$\begin{aligned}
 \langle \phi_b | U(+\infty, -\infty) | \phi_a \rangle &= \langle \psi_b^- | \psi_a^+ \rangle = \left[\langle \phi_b | + \langle \phi_b | V \frac{1}{E_b + i\varepsilon - H} \right] | \psi_a^+ \rangle \\
 &= \langle \phi_b | \phi_a \rangle + \langle \phi_b | \frac{1}{E_a + i\varepsilon - H_0} V | \psi_a^+ \rangle + \langle \phi_b | V \frac{1}{E_b + i\varepsilon - H} | \psi_a^+ \rangle \\
 &= \langle \phi_b | \phi_a \rangle - 2\pi i \delta_\varepsilon(E_b - E_a) \langle \phi_b | V | \psi_a^+ \rangle
 \end{aligned} \tag{12.9}$$

insofar

$$| \psi_b^- \rangle = U(0, +\infty) | \phi_b \rangle, \quad | \psi_a^+ \rangle = U(0, -\infty) | \phi_a \rangle,$$

and from eq.(4.4) $U^+(0, +\infty) = U(+\infty, 0)$. The S-matrix will be defined by

$$\langle \phi_b | S | \phi_a \rangle = \langle \phi_b | U(+\infty, -\infty) | \phi_a \rangle = \langle \psi_b^- | \psi_a^+ \rangle. \tag{12.10}$$

12.5 The Møller wave function operators and the S-matrix

Let us introduce two operators, the “Møller wave function operators”

$$\Omega_+ = U(0, -\infty), \quad \Omega_- = U(0, +\infty).$$

We have

$$| \psi_a^+ \rangle = \Omega_+ | \phi_a \rangle.$$

In potential scattering $| \psi_a^+ \rangle$ is represented by a plane wave plus an outgoing spherical wave and

$$| \psi_a^- \rangle = \Omega_- | \phi_a \rangle$$

is represented by a plane wave plus an incoming spherical wave.

The operator Ω_+ is isometric i.e.

$$\Omega_+^\dagger \Omega_+ = I \tag{12.11}$$

In fact from the relations (12.5) we have

$$\langle \phi_b | \Omega_+^\dagger \Omega_+ | \phi_a \rangle = \langle \psi_b^+ | \psi_a^+ \rangle = \langle \phi_b | \phi_a \rangle$$

from which eq.(12.11) follows, being $| \phi_a \rangle$ a complete set. Similarly one proves that Ω_- is isometric. Due to isometry we have that $R_+ = \text{Range } \Omega_+$ is a subspace and

the same holds true for $R_- = \text{Range } \Omega_-$. It is an assumption which characterizes a scattering system [3] that such two subspaces coincide i.e. $R_+ = R_- \equiv R$. The physical meaning of this statement is that given a state at time zero, which for large negative times evolves according to H_0 the same state at large positive times evolves according to H_0 and vice-versa. For a more general treatment see [6]. The orthogonal complement of R will be denoted by Q .

Then it follows that

$$\Omega_+ \Omega_+^\dagger = I - B$$

being B the projector on the subspace Q . In fact if $f \in R$ we have that there exist a g such that $f = \Omega_+ g$ and then $\Omega_+ \Omega_+^\dagger f = \Omega_+ g = f$. If on the other hand $h \in Q$ we have $(\Omega_+ g, h) = 0$ for all g and thus $\Omega_+^\dagger h = 0$.

Similarly

$$\Omega_-^\dagger \Omega_- = I$$

$$\Omega_- \Omega_-^\dagger = I - B .$$

We examine now the unitarity of the S -matrix. From (12.10) we have $S = \Omega_-^\dagger \Omega_+$ and

$$S^\dagger S = \Omega_+^\dagger \Omega_- \Omega_-^\dagger \Omega_+ = \Omega_+^\dagger (I - B) \Omega_+ = I - \Omega_+^\dagger B \Omega_+ = I$$

$$SS^\dagger = \Omega_-^\dagger \Omega_+ \Omega_+^\dagger \Omega_- = \Omega_-^\dagger (I - B) \Omega_- = I - \Omega_-^\dagger B \Omega_- = I .$$

The fact that $\Omega_+^\dagger B \Omega_+$ equals zero is proved by applying such an operator to the generic vector $|\phi_a\rangle$. In fact

$$\Omega_+ |\phi_a\rangle = |\psi_a^+\rangle$$

and

$$B |\psi_a^+\rangle = 0 .$$

12.6 The unitarity of the S -matrix and the optical theorem

We can express the unitarity of the S -matrix in terms of the matrix $\mathcal{T}$ defined by the relation

$$S = I - i\mathcal{T}$$

We have

$$S^+ S = I + i(\mathcal{T}^+ - T) + \mathcal{T}^+ \mathcal{T} = I$$

from which

$$i(\mathcal{T} - \mathcal{T}^+) = \mathcal{T}^+ \mathcal{T} .$$

We define T by

$$\langle \phi_b | \mathcal{T} | \phi_a \rangle = 2\pi \delta(E_b - E_a) \langle \phi_b | T | \phi_a \rangle .$$

From eq.(12.9) we have

$$\langle \phi_b | T | \phi_a \rangle = \langle \phi_b | V | \psi_a^+ \rangle . \quad (12.12)$$

Simplifying the notation we obtain

$$\langle b | \mathcal{T} | a \rangle - \langle b | \mathcal{T}^+ | a \rangle = -i \int \langle b | \mathcal{T}^+ | n \rangle dn \langle n | \mathcal{T} | a \rangle$$

i.e.

$$2\pi i \delta(E_b - E_a) (\langle b | T | a \rangle - \langle b | T^+ | a \rangle) = (2\pi)^2 \int \langle b | T^+ | n \rangle dn \langle n | T | a \rangle \delta(E_a - E_b) \delta(E_a - E_n)$$

with

$$\int |n\rangle dn \langle n| = I$$

i.e.

$$i [\langle b | T | a \rangle - (\langle a | T | b \rangle)^*] = 2\pi \int (\langle n | T | b \rangle)^* \langle n | T | a \rangle \delta(E_a - E_n) dn . \quad (12.13)$$

For $|a\rangle = |b\rangle$ we obtain the optical theorem

$$-2 \operatorname{Im} \langle a | T | a \rangle = 2\pi \int |\langle n | T | a \rangle|^2 \delta(E_n - E_a) dn .$$

If T is symmetric under the exchange of $|a\rangle$ with $|b\rangle$ the first term in eq.(12.13) will be equal to $-2 \operatorname{Im} \langle b | T | a \rangle$. This occurs in a very important case i.e. when the theory is invariant under time reversal and parity and when we adopt for the vectors $|a\rangle$ and $|b\rangle$ special phase conventions. We want to see how this happens in the simple case of non relativistic potential scattering. To say that the theory is invariant under time reversal means that (in absence of spin)

$$i \frac{\partial \psi}{\partial t} = H \psi$$

implies

$$-i\frac{\partial\psi^*}{\partial t} = H\psi^*$$

where we have $H = H_0 + V$. When spin is present see [5]. Clearly a local and thus real potential gives rise to an hamiltonian invariant under time reversal. A non local potential $V(\mathbf{x}, \mathbf{x}')$ can give a counterexample. In order for V to be hermitean we need $V(\mathbf{x}, \mathbf{x}') = V^*(\mathbf{x}', \mathbf{x})$. To be the hamiltonian invariant under time reversal we need

$$\left(\int V(\mathbf{x}, \mathbf{x}')\psi(\mathbf{x}')d\mathbf{x}' \right)^* = \int V(\mathbf{x}, \mathbf{x}')\psi^*(\mathbf{x}')d\mathbf{x}'$$

i.e. V must be real and thus symmetric. Let us examine the consequences of such a potential on the related T -matrix.

$$\langle\phi_b|S|\phi_a\rangle = \langle\phi_b|\phi_a\rangle - 2\pi i\delta(E_b - E_a)\langle\phi_b|T|\phi_a\rangle .$$

From eq.(12.12)

$$\langle\phi_b|T|\phi_a\rangle = \langle\phi_b|V|\psi_a^+\rangle = \langle\phi_b|V|\phi_a\rangle + \langle\phi_b|V\frac{1}{E_a - H_0 + i\varepsilon}T|\phi_a\rangle$$

and also

$$\langle\phi_b|T|\phi_a\rangle = \langle\phi_b|V|\psi_a^+\rangle = \langle\phi_b|V|\phi_a\rangle + \langle\phi_b|V\frac{1}{E_a - H + i\varepsilon}V|\phi_a\rangle$$

which we can write as

$$\begin{aligned} T &= V + V\frac{1}{E - H_0 + i\varepsilon}T , \\ T &= V + V\frac{1}{E - H + i\varepsilon}V . \end{aligned} \tag{12.14}$$

Let us consider now the symmetry properties of the matrix elements of V

$$\begin{aligned} \langle\phi_b|V|\phi_a\rangle &= \frac{1}{(2\pi)^3} \int e^{-i\mathbf{k}_b \cdot \mathbf{x}} V(\mathbf{x}, \mathbf{x}') e^{i\mathbf{k}_a \cdot \mathbf{x}'} d\mathbf{x} d\mathbf{x}' \\ &= \frac{1}{(2\pi)^3} \int e^{i\mathbf{k}_a \cdot \mathbf{x}} V(\mathbf{x}, \mathbf{x}') e^{-i\mathbf{k}_b \cdot \mathbf{x}'} d\mathbf{x} d\mathbf{x}' \end{aligned}$$

i.e.

$$\langle\mathbf{k}|V|\mathbf{k}'\rangle = \langle-\mathbf{k}'|V|-\mathbf{k}\rangle$$

in the representation in which

$$\langle \mathbf{x} | \mathbf{k} \rangle = \frac{e^{i\mathbf{k} \cdot \mathbf{x}}}{(2\pi)^{3/2}} .$$

For the kinetic energy we have

$$\langle \mathbf{k} | \mathbf{p}^2 | \mathbf{k}' \rangle = \mathbf{k}^2 \delta(\mathbf{k} - \mathbf{k}') = \langle -\mathbf{k}' | \mathbf{p}^2 | -\mathbf{k} \rangle$$

and as a consequence we have from (12.14)

$$\begin{aligned} \langle \mathbf{k} | T | \mathbf{k}' \rangle &= \langle \mathbf{k} | V | \mathbf{k}' \rangle + \int \int \langle \mathbf{k} | V | \mathbf{h} \rangle d\mathbf{h} \langle \mathbf{h} | \frac{1}{E - H + i\varepsilon} | \mathbf{h}' \rangle d\mathbf{h}' \langle \mathbf{h}' | V | \mathbf{k}' \rangle \\ \langle -\mathbf{k}' | T | -\mathbf{k} \rangle &= \langle \mathbf{k} | V | \mathbf{k}' \rangle + \int \int \langle -\mathbf{k}' | V | \mathbf{h} \rangle d\mathbf{h} \langle \mathbf{h} | \frac{1}{E - H + i\varepsilon} | \mathbf{h}' \rangle d\mathbf{h}' \langle \mathbf{h}' | V | -\mathbf{k} \rangle \\ &= \langle \mathbf{k} | V | \mathbf{k}' \rangle + \int \int \langle \mathbf{k} | V | -\mathbf{h} \rangle d\mathbf{h} \langle -\mathbf{h} | \frac{1}{E - H + i\varepsilon} | -\mathbf{h}' \rangle d\mathbf{h}' \langle -\mathbf{h}' | V | \mathbf{k}' \rangle = \langle \mathbf{k} | T | \mathbf{k}' \rangle . \end{aligned}$$

i.e.

$$\langle \mathbf{k} | T | \mathbf{k}' \rangle = \langle -\mathbf{k}' | T | -\mathbf{k} \rangle .$$

The conclusion is that the events depicted in Figure 12.2 have the same ampli-

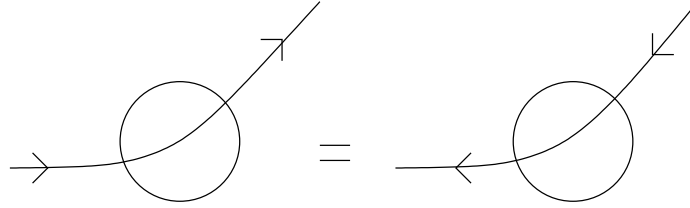


Figure 12.2: Time reversed process

tude. Moreover if the theory is invariant under parity

$$\langle \mathbf{k} | V | \mathbf{k}' \rangle = \langle -\mathbf{k} | V | -\mathbf{k}' \rangle = \langle \mathbf{k}' | V | \mathbf{k} \rangle$$

and it follows

$$\langle \mathbf{k} | T | \mathbf{k}' \rangle = \langle \mathbf{k}' | T | \mathbf{k} \rangle . \quad (12.15)$$

However the simple form of eq.(12.15) depends in essential way on the used representation. If e.g. we change over to

$$|\mathbf{k}\rangle \rightarrow e^{i\alpha(\mathbf{k})}|\mathbf{k}\rangle = |\mathbf{k}\rangle\rangle$$

we have

$$\langle\langle\mathbf{k}'|V|\mathbf{k}\rangle\rangle = \langle\langle\mathbf{k}|V|\mathbf{k}'\rangle\rangle e^{2i(\alpha(\mathbf{k})-\alpha(\mathbf{k}'))} \neq \langle\langle\mathbf{k}|V|\mathbf{k}'\rangle\rangle .$$

12.7 Transition probabilities per unit time

The expression

$$w = |\langle\phi_b|U(t, -\infty)|\phi_a\rangle|^2$$

with $b \neq a$ gives us the probability to find at time t the particle with momentum $\mathbf{p}_b$. Being the process stationary and recalling that t represents the time divided by $\hbar$, the transition probability per unit time to the state of momentum $\mathbf{p}_b$ using (12.1) and (12.4) is given by

$$\begin{aligned} \frac{1}{\hbar} \frac{\partial w}{\partial t} &= \frac{1}{\hbar} \langle\phi_b|U(0, -\infty)|\phi_a\rangle^* \langle\phi_b|\frac{\partial U(t, -\infty)}{\partial t}|\phi_a\rangle|_{t=0} + \text{c.c.} \\ &= -\frac{i}{\hbar} \langle\phi_b|V|\psi_a^+\rangle^* \frac{1}{E_a - E_b - i\varepsilon} \langle\phi_b|V|\psi_a^+\rangle + \text{c.c.} \end{aligned}$$

i.e.

$$\frac{1}{\hbar} \frac{\partial w}{\partial t} = \frac{2\pi}{\hbar} \delta(E_b - E_a) |\langle\phi_b|V|\psi_a^+\rangle|^2 = \frac{2\pi}{\hbar} \delta(E_b - E_a) |\langle\phi_b|T|\phi_a\rangle|^2 . \quad (12.16)$$

From this expression one goes over to the cross section with standard methods. The differential cross section is defined as the transition probability from an initial state ϕ_a , to all final states whose versor $\mathbf{k}/k$ lies in a given solid angle $\Delta\Omega$ and dividing by the initial flux. For

$$\langle\mathbf{x}|\phi_a\rangle = \frac{1}{\sqrt{(2\pi)^3}} e^{i\mathbf{k}_a \cdot \mathbf{x}}$$

we have the initial flux

$$\frac{\hbar k_a}{(2\pi)^3 M} .$$

Integrating (12.16) in $d\mathbf{k}$ with $\mathbf{k}/k$ lying in $\Delta\Omega$ and dividing by the initial flux we have

$$\frac{d\sigma(\theta, \phi)}{d\Omega} \Delta\Omega = \left(\frac{2\pi}{\hbar}\right)^4 M^2 |\langle \phi_b | V | \psi_a^+ \rangle|^2 \Delta\Omega = \left(\frac{2\pi}{\hbar}\right)^4 M^2 |\langle \phi_b | T | \phi_a \rangle|^2 \Delta\Omega$$

in agreement with eq.(12.7).

References

- [1] B.A. Lippmann and J. Schwinger *Variational principles for scattering processes I*, Phys.Rev. **79** 3 (1950)
- [2] M. Gell-Mann and M.L. Goldberger *The formal theory of scattering*, Phys.Rev. **91** 2 (1953)
- [3] J.M. Jauch *Theory of the scattering operator*, Helv.Phys.Acta **31** 127 (1958)
- [4] E.T. Whittaker and G.N. Watson *A course of modern analysis*, Chap.VI, Cambridge University Press, Cambridge U.K. (2010)
- [5] J. Hamilton *The theory of elementary particles*, Oxford, Clarendon press (1959)
- [6] M. Reed, B. Simon *Methods of modern mathematical physics III: scattering theory*, Academic Press, New York (1979)

Chapter 13

Scattering by a central potential

13.1 Introduction

In this chapter we shall examine in detail some properties of the function $S_0(k)$ i.e. of the S -matrix in the partial wave $l = 0$, in the model of non relativistic potential scattering.

The reduced radial wave function obeys the Schrödinger equation

$$-\varphi'' + U(r)\varphi = k^2\varphi, \quad \text{with} \quad U = \frac{2mV}{\hbar^2}.$$

We shall confine ourselves to potentials which at the origin do not diverge more rapidly than $\frac{1}{r^{2-\varepsilon}}$, $\varepsilon > 0$. In this case from the indicial equation one deduces that one of the possible solution vanishes for $r \rightarrow 0$ and has derivative $\varphi'(0) \neq 0$ and finite. The condition $\varphi(0) = 0$ chooses among the solutions the physically acceptable one (see Chapter V). We shall suppose furthermore that our potential vanishes sufficiently fast (faster than $r^{-1-\varepsilon}$) for $r \rightarrow \infty$. The solution will have the asymptotic behavior

$$\varphi \sim e^{-ikr} - S(k)e^{ikr} \tag{13.1}$$

where the coefficient $S(k)$ is the element of the S -matrix insofar for large r

$$e^{-ikr} - S(k)e^{ikr} = -2ie^{i\delta} \sin(kr + \delta), \quad S(k) = e^{2i\delta}.$$

S is completely determined by the condition $\varphi(0) = 0$. Let us suppose that the solutions of the Schrödinger equation are meaningful also for complex k . This will be rigorously shown in the following.

Then the function S will be defined also for non physical values of k . For this $S(k)$ some general properties hold. For example for $k \rightarrow k^*$ the solution will be φ^* , being U real.

$$\varphi^*(k, r) \sim (e^{-ikr} - S(k)e^{ikr})^* = e^{ik^*r} - S(k)^*e^{-ik^*r}$$

from which it follows

$$S^*(k) = \frac{1}{S(k^*)} .$$

For $k \rightarrow -k$ we have

$$\varphi(-k, r) = e^{ikr} - S(-k)e^{-ikr}$$

from which it follows

$$S(-k) = \frac{1}{S(k)} . \quad (13.2)$$

Thus

$$S^*(k) = \frac{1}{S(k^*)} = S(-k^*) . \quad (13.3)$$

Notice that while (13.2) is general (13.3) is related to the fact that being U local, U is real. Such a property is related to the invariance of the theory under time reversal (see Chapter 12). Thus we have the scheme of Fig.(13.1) .

For pure imaginary values of k , Schrödinger equation can describe bound states. In fact let $k = i\chi$; to have a bound state of energy $-\chi^2$ it is necessary that the wave function φ , vanishing at the origin behaves at infinity like $e^{-\chi r}$. I.e.

$$e^{\chi r} - S(i\chi)e^{-\chi r} = ce^{-\chi r}, \quad c = \text{const}$$

which means that $S(k)$ must have a pole for $k = i\chi$ and thus a zero for $k = -i\chi$ as $S^{-1}(k) = S(-k)$. For a pole for negative χ we obtain a solution which explodes at infinity as a pure exponential. These last solutions take the name of anti-bound states or virtual states; their occurrence gives rise to physically observable effects even though they do not correspond to physical states.

To solutions which vanish at the origin and behave as pure complex exponentials at infinity there correspond poles of the S matrix. Such poles with real part

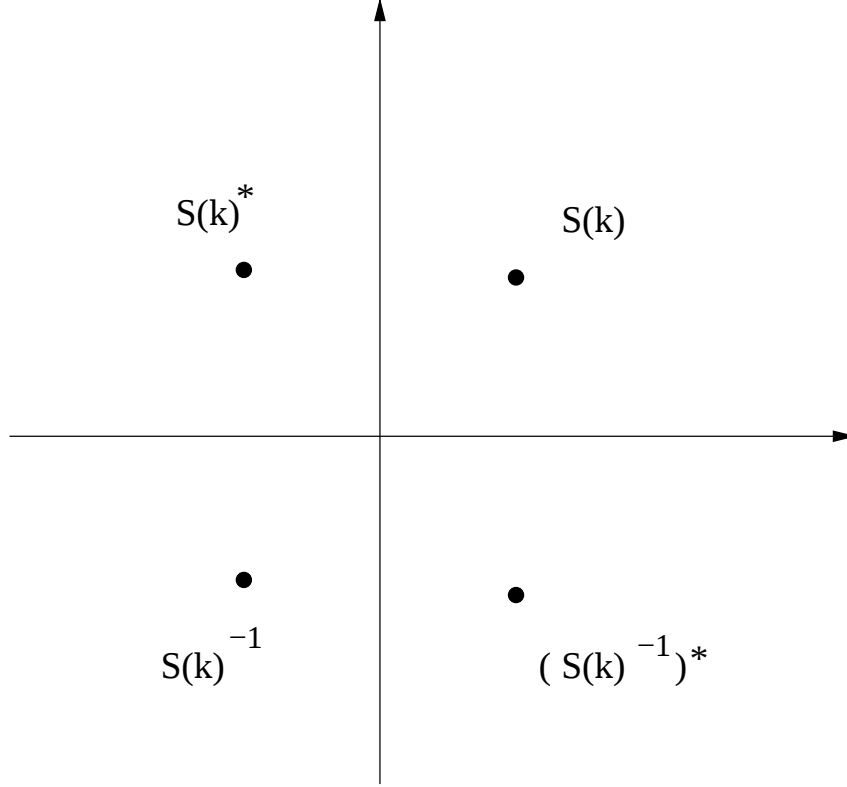


Figure 13.1: Inversion and conjugation

different from zero can exist only in the lower half plane. In fact let φ be a solution corresponding to a pole of S in $k = R + iI$, with $I > 0$. We have

$$\begin{aligned} -\varphi'' + U\varphi &= k^2\varphi \\ -\varphi^{*''} + U\varphi^* &= (k^2)^*\varphi^* \end{aligned}$$

from which

$$-\frac{\partial}{\partial r}(\varphi'\varphi^* - \varphi^{*'}\varphi) = (k^2 - k^{*2})\varphi^*\varphi. \quad (13.4)$$

Being $I > 0$ $\varphi^*\varphi$ goes to zero exponentially at infinity, $|\varphi| \sim e^{-r\text{Im}k}$ and $\varphi(0) = 0$. Integrating (13.4) we have

$$(k^2 - k^{*2}) \int_0^\infty \varphi^*\varphi dr = [\varphi^*\varphi' - \varphi^{*'}\varphi]_0^\infty = 0 \quad (13.5)$$

so that $k^2 - k^{*2} = 0$, i.e. $R = 0$. The behavior of S for real k , $k \rightarrow \infty$ can be investigated from the exact relation

$$4ke^{i\delta} \sin \delta = \int_0^\infty U(r) \varphi(r) \varphi_0(r) dr \quad (13.6)$$

which is easily obtained from the Schrödinger equation. In (13.6) φ is given by the solution with the asymptotic behavior (13.1) and $\varphi_0(r) = e^{-ikr} - e^{ikr}$.

In the Born approximation it gives

$$k \sin \delta = - \int_0^\infty U(r) \sin^2 kr \, dr$$

and thus for $k \rightarrow \infty$ and potentials whose integral converge at zero and at infinity

$$\delta(k) \approx -\frac{1}{2k} \int_0^\infty U(r) dr$$

and we shall have $\delta(\infty) = 0$ and $S(\infty) = 1$. Similar considerations can be carried out also for certain potentials whose integral in r does not converge. The distributions of poles ($\times$) and of zeros ($\circ$) is of the type given in Fig.(13.2).

13.2 The Jost functions

We can deal systematically with the problems of the preceding paragraph introducing some particular solutions of the Schrödinger equation

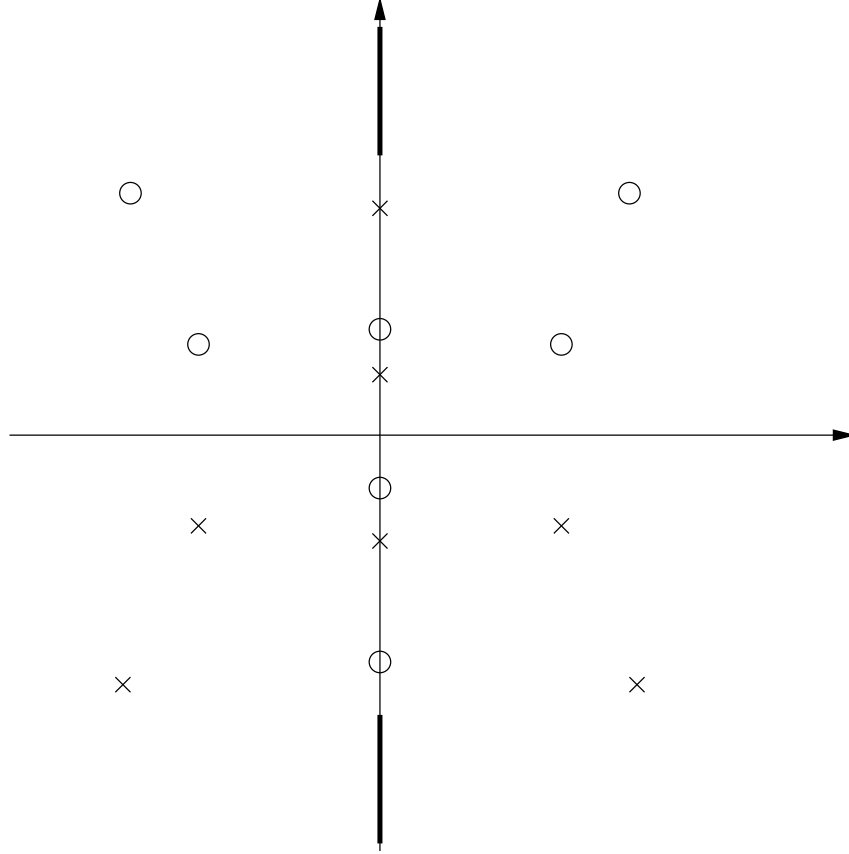
$$-f''(k, r) + U(r)f(k, r) = k^2 f(k, r)$$

with asymptotic behavior $f(k, r) \sim e^{-ikr}$ for $r \rightarrow \infty$ and as such do not necessarily satisfy the boundary condition $f(k, 0) = 0$. Such solutions are called the Jost functions [1]. In addition to $f(k, r)$ also $f(-k, r)$ is a solution which is independent of $f(k, r)$: its Wronskian with $f(k, r)$ is $-2ik$. The regular solution $\varphi(k, r)$ with $\varphi(k, 0) = 0$ and $\varphi'(k, 0) = 1$ is linearly dependent on $f(k, r)$ and $f(-k, r)$ i.e.

$$\varphi(k, r) = C f(k, r) + D f(-k, r) . \quad (13.7)$$

To compute C and D we consider the Wronskian computed at $r = 0$

$$w(\varphi, f) = \varphi'(k, 0)f(k, 0) - \varphi(k, 0)f'(k, 0) = f(k, 0)$$

Figure 13.2: Singularities of the S -matrix

and that derived from (13.7)

$$w(\varphi, f) = C w(f(k, r), f(k, r)) + D w(f(-k, r), f(k, r)) = 2ikD$$

and thus

$$D = \frac{f(k, 0)}{2ik}, \quad \text{and} \quad C = -\frac{f(-k, 0)}{2ik}.$$

In the following we use the notation $f(k) = f(k, 0)$. Thus

$$\varphi(k, r) = \frac{1}{2ik} [f(k)f(-k, r) - f(-k)f(k, r)]$$

and being for large r

$$\varphi(k, r) \approx \frac{1}{2ik} [f(k)e^{ikr} - f(-k)e^{-ikr}]$$

we shall have

$$S(k) = \frac{f(k)}{f(-k)} .$$

From the definition of $f(k, r)$ and from the uniqueness of the solutions it follows that $f(-k^*, r) = [f(k, r)]^*$ and thus $f(-k^*) = [f(k)]^*$ and from these the symmetry properties (13.2, 13.3) of $S(k)$. We study now the mathematical properties of $f(k, r)$ [2]. It is useful to transform the differential equation into an integral equation, which includes the conditions on the asymptotic behavior. The following Volterra equation is particularly apt

$$f(k, r) = e^{-ikr} + \frac{1}{k} \int_r^\infty U(x) \sin k(x-r) f(k, x) dx . \quad (13.8)$$

A derivation is the following

$$f'' + k^2 f = U f$$

$$\frac{d^2}{dr^2} \sin k(x-r) + k^2 \sin k(x-r) = 0$$

$$\begin{aligned} U(x) f(k, x) \sin k(x-r) &= f''(k, x) \sin k(x-r) - f(k, x) \frac{d^2}{dx^2} \sin k(x-r) \\ &= \frac{d}{dx} [f'(k, x) \sin k(x-r) - f(k, x) \frac{d}{dx} \sin k(x-r)] \end{aligned}$$

and integrating we have

$$\int_r^\infty U(x) \sin k(x-r) f(k, x) dx = -k e^{-ikr} + k f(k, r) .$$

Rewriting now (13.8) with $f(k, r) = e^{-ikr} g(k, r)$ we obtain

$$g(k, r) = 1 + \frac{1}{k} \int_r^\infty U(x) \frac{1 - e^{-2ik(x-r)}}{2i} g(k, x) dx . \quad (13.9)$$

One can solve such an equation by iteration

$$\begin{aligned} g(k, r) &= \sum_{n=0}^{\infty} g_n(k, r) \\ g_0(k, r) &= 1 \\ g_1(k, r) &= \frac{1}{k} \int_r^\infty U(x) \frac{1 - e^{-2ik(x-r)}}{2i} dx \end{aligned}$$

$$g_{n+1}(k, r) = \frac{1}{k} \int_r^\infty U(x) \frac{1 - e^{-2ik(x-r)}}{2i} g_n(k, x) dx .$$

Let us analyze now the conditions for the convergence of the above series. We take k complex and we shall distinguish two cases 1) $\text{Im}k = b \leq 0$; 2) $\text{Im}k = b > 0$.

1) $\text{Im}k = b \leq 0$ with $k \neq 0$.

We have

$$\left| \frac{1 - e^{-2ik(x-r)}}{2ik} \right| \leq \left| \frac{1}{2ik} \right| + \left| \frac{e^{2b(x-r)}}{2ik} \right| \leq \frac{1}{|k|}$$

from which

$$|g_{n+1}(k, r)| \leq \frac{1}{|k|} \int_r^\infty |U(x)| |g_n(k, x)| dx$$

and thus

$$|g_1(k, r)| \leq \frac{1}{|k|} \int_r^\infty |U(x)| dx \equiv \frac{1}{|k|} M(r) \quad (13.10)$$

$$|g_2(k, r)| \leq \frac{1}{|k|^2} \int_r^\infty |U(x)| M(x) dx = -\frac{1}{|k|^2} \int_r^\infty \frac{dM(x)}{dx} M(x) dx = \frac{M(r)^2}{2|k|^2}$$

and by induction

$$|g_n(k, r)| \leq \frac{M^n(r)}{n!|k|^n}$$

$$|g(k, r) - 1| \leq e^{\frac{M(r)}{|k|}} - 1 .$$

Thus we have:

a) If $b \leq 0$, $k \neq 0$ and $\int_r^\infty |U(x)| dx < \infty$ for $r > 0$, then the series is a solution of (13.9) as it converges uniformly.

b) Each term of the series is analytic in k .

c) If $b < 0$, ($k \neq 0$) the series converges on a closed contour. Then by a well known theorem the series is analytic in k .

d) For $k \rightarrow \infty$ on the real axis and in the lower half-plane we have $g(r, k) \rightarrow 1$

e) If $M(0) < \infty$ all the statements holds also for $g(k, 0) = f(k, 0)$.

A sufficient condition for having $M(0) < \infty$ is that the potential at infinity be bounded by $r^{-1-\varepsilon}$ and at the origin by $r^{-1+\varepsilon}$.

A different way to proceed, which includes $k = 0$ in the convergence region of the series, is to use the inequality

$$\left| \frac{1 - e^{-2ik(x-r)}}{2ik} \right| = \left| \int_0^{x-r} e^{-2i\eta k} d\eta \right| \leq x - r < x .$$

The changes in the proof are

- a) k is absorbed in the majoring factor
- b) $U(x)$ is multiplied by x . Thus we have

$$|g(k, r) - 1| \leq e^{N(r)} - 1 \quad \text{with} \quad N(r) = \int_r^\infty x|U(x)|dx$$

which requires a faster convergence of $U(x)$ at infinity.

2) $\text{Im}k = b > 0$ ($k \neq 0$)

$$\left| \frac{1 - e^{-2ik(x-r)}}{2ik} \right| \leq \frac{e^{2b(x-r)}}{|k|} \quad \text{because} \quad e^{2b(x-r)} \geq 1$$

$$|g_1(k, r)| \leq \int_r^\infty |U(x)| \frac{e^{2b(x-r)}}{|k|} dx \leq \frac{e^{-2br}}{|k|} P(r)$$

with

$$P(r) = \int_r^\infty |U(x)| e^{2bx} dx$$

$$|g_2(k, r)| \leq \int_r^\infty |U(x)| \frac{e^{2b(x-r)}}{|k|} \frac{e^{-2bx}}{|k|} P(x) dx \leq \frac{e^{-2br}}{|k|^2} \int_r^\infty |U(x)| P(x) dx \leq P(r) \frac{e^{-2br}}{|k|^2} M(r)$$

with $M(r)$ given in eq(13.10). Then

$$|g_n(k, r)| \leq P(r) \frac{e^{-2br}}{|k|^n} \frac{M^{n-1}(r)}{(n-1)!}$$

$$|g(k, r) - 1| \leq P(r) \frac{e^{-2br}}{|k|} e^{\frac{M(r)}{|k|}}.$$

We also have

$$\left| \frac{1 - e^{-2ik(x-r)}}{2ik} \right| = \left| \int_0^{x-r} e^{-2ik\eta} d\eta \right| \leq (x-r) e^{2b(x-r)} \leq x e^{2b(x-r)}.$$

Then defined $Q(r)$ by

$$|g_1(k, r)| \leq \int_r^\infty |U(x)| x e^{2b(x-r)} |g_0(x, k)| dx = e^{-2br} \int_r^\infty |U(x)| x e^{2bx} dx \equiv e^{-2br} Q(r)$$

we have

$$\begin{aligned} |g_2(k, r)| &\leq e^{-2br} \int_r^\infty |U(x)|x dx \int_x^\infty |U(y)|y e^{2by} dy \leq \\ &e^{-2br} \int_r^\infty |U(y)|y e^{2by} dy \int_r^\infty |U(x)|x dx \leq e^{-2br} Q(r)N(r) \end{aligned}$$

with

$$N(r) = \int_r^\infty x|U(x)|dx, \quad Q(r) = \int_r^\infty x|U(x)|e^{2bx}dx .$$

In this way we reach the result

$$|g_n(k, r)| \leq Q(r) \frac{N^{n-1}(r)}{(n-1)!} e^{-2br}$$

and

$$|g(k, r) - 1| \leq Q(r) e^{-2br} e^{N(r)} .$$

Then the Jost function $f(k, r)$ is an analytic function of k for fixed r for $\text{Im } k < \mu/2$ being $\mu/2$ the upper limit of the values of b for which

$$\int_r^\infty x|U(x)|e^{2bx}dx < \infty .$$

For $r \neq 0$ we have that for $|k| \rightarrow \infty$ the function $g(k, r)$ goes to 1 in the region $\text{Im } k < \frac{\mu}{2}$. Such results extend to $r = 0$ i.e. to $f(k) = f(k, 0)$ when $Q(0) < \infty$.

Thus we have shown that the $S(k) = f(k)/f(-k)$ is a meromorphic function in the strip $|\text{Im } k| < \mu/2$, called the Bargmann strip [3]. The poles of $S(k)$ in the interior of this strip do not arise from singularities of $f(k)$ but from zeros of $f(-k)$. The argument given at eqs.(13.4,13.5) shows that in the lower half plane $f(k)$ can have zeros only on the imaginary axis and that to those zero there correspond the bound states. Their wave function is given by

$$\varphi(k, r) = \frac{f(k)f(-k, r)}{2ik} \sim \frac{f(k)e^{ikr}}{2ik} = -\frac{f(i\chi)e^{-\chi r}}{2\chi} .$$

We point out that the zeros of $f(k)$ in the lower half-plane are simple zeros. In fact if we take the derivative with respect to k of the Schrödinger equation for the solution

$$\phi(k, r) = f(-k)f(k, r) - f(k)f(-k, r)$$

multiply by $\phi(r)$ and subtract from the result the Schrödinger equation multiplied by $\partial\phi/\partial k$ and integrate we obtain

$$\left[-\phi \frac{\partial\phi'}{\partial k} + \phi' \frac{\partial\phi}{\partial k} \right]_0^r = 2k \int_0^r \phi^2 dr \quad (13.11)$$

For $k = -i\chi$ we have the real bound state solution

$$\phi_B = f(i\chi)f(-i\chi, r)$$

and eq.(13.11) gives

$$if(i\chi) \frac{df}{dk}(-i\chi) = \int_0^\infty \phi_B^2 dr .$$

and we obtain $\frac{df}{dk}(-i\chi) \neq 0$.

The Bargmann strip extends to the whole k -plane for the so called finite-range potentials; by these we understand those potentials which decrease at infinity faster than any exponential, e.g. a gaussian potential. In these cases b is arbitrary. The analyticity region of $S(k)$ can be notably extended if the potential in question is yukawian i.e. the superposition discrete or continuous of Yukawa potentials

$$U(r) = \int_m^\infty \frac{e^{-\mu r}}{r} \sigma(\mu) d\mu .$$

Several condition can be imposed on $\sigma(\mu)$ e.g

$$\int_m^\infty |\sigma(\mu)| d\mu < \infty .$$

In this case the domain of analyticity of $f(k)$ is extended to the whole complex plane except for the cut $(i m/2, +i\infty)$.

The structure of the $S(k)$ matrix for yukawian potentials is shown in Figure 13.3. The cut is not necessarily of continuous structure. There exist potentials like e.g. the Hultèn potential

$$U(r) = \lambda \frac{e^{-\mu r}}{1 - e^{-\mu r}}$$

for which such a cut reduces to an infinite number of poles.

One notices that the distance of the cut from the origin $im/2$ is directly related to the “range” $1/m$ of the potential in question.

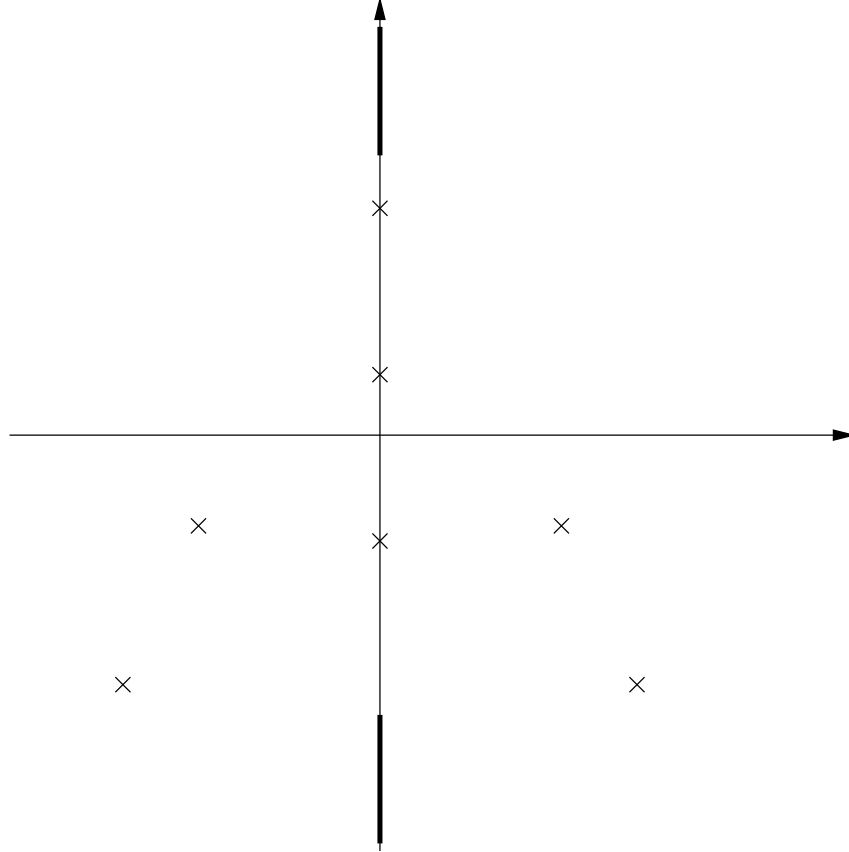


Figure 13.3: Singularity structure of the S -matrix for yukawian potentials

13.3 Physical interpretation of the singularities

Let us examine which is the physical meaning of the position and of the residues of the poles of the S -matrix. We already saw that to a bound state there correspond a pole on the imaginary axis in the upper half plane whose position is related in simple way to the binding energy of the bound state in question. As far as the residue is concerned it is useful to exploit the relation $S(k) = 1/S(-k)$ and examine the zeros of S in the lower half plane. To compute the residue at the pole we repeat the same reasoning which led us to eq.(13.11) using the wave function

$$\varphi(k, r) = f(k, r) - S(k)f(-k, r) .$$

We obtain at the point $k = -i\chi$ where $S(k)$ vanishes

$$\frac{\partial S(-i\chi_B)}{\partial k} = -i \int_0^\infty \varphi_B^2(r) dr$$

being φ_B the real wave function of the bound state with asymptotic behavior $\varphi_B(r) \approx e^{-\chi_B r}$. In the neighborhood of the pole then we have

$$S(k) = \frac{1}{S(-k)} \approx \frac{-i\alpha}{k - i\chi_B}$$

with

$$\alpha^{-1} = \int_0^\infty \varphi_B^2(x) dx .$$

The fact that the residue has to be pure imaginary is seen also from $S(k) = S^*(-k^*)$. In the variable k^2 we have with $S(k) = S_2(k^2)$

$$S_2(k^2) \sim \frac{-i\alpha 2i\chi_B}{(k - i\chi_B)(k + i\chi_B)} = \frac{2\alpha\chi_B}{k^2 + \chi_B^2} .$$

We can now relate the residue α with the asymptotic normalization of the wave function. If

$$\int \varphi^2(r) dr = 1 \quad \text{and} \quad \beta \varphi_B(r) = \varphi(r) \approx \beta e^{-\chi_0 r}$$

$$\int \varphi_B^2(r) dr = \frac{1}{\alpha} = \frac{1}{\beta^2}$$

$$S(k) \approx \frac{-i\beta^2}{k - i\chi_B} .$$

If the pole is near the origin with respect to the other singularities we can use the approximation

$$f(k) \approx ic(k + i\chi_B)$$

where c is real due to $f(-k^*) = [f(k)]^*$. It follows that

$$S(k) \approx -\frac{k + i\chi_B}{k - i\chi_B}$$

from which $\beta^2 \approx 2\chi_B$ and thus for $\chi_B \rightarrow 0$ we have $\beta^2 \rightarrow 0$. It is possible to relate these results with the cross section at low energies.

$$e^{2i\delta} = \frac{\chi_B - ik}{\chi_B + ik}$$

from which

$$\tan \delta = -\frac{k}{\chi_B} .$$

In our case $\delta \sim -k/\chi_B$ and recalling that for small k , $\delta = -ak$ with a the scattering length we have

$$\sigma_s(0) = 4\pi a^2 = \frac{4\pi}{\chi_B^2} .$$

This is the case for the deuteron in the $p-n$ scattering.

A zero of the Jost function $f(k)$ on the upper imaginary axis is called an antibound state. In this case the (unphysical) Schrödinger wave function behaves at infinity as $e^{\chi_A r}$.

In the case of an antibound state we have the situation

$$S(k) \approx -\frac{k - i\chi_A}{k + i\chi_A}$$

and thus $\delta \approx k/\chi_A$ and

$$\sigma_s(0) = 4\pi a^2 = \frac{4\pi}{\chi_A^2} .$$

One sees that there is no difference in the cross section but the sign of the scattering length is opposite.

Let us consider now a pole of the S -matrix in the lower half plane of position $k_R - i\Gamma$ with $0 < \Gamma \ll k_R$. To this pole there corresponds a resonance. The Jost function in the proximity of such a point will have the form $f(k) \approx c(k - (k_R + i\Gamma))$ from which

$$S = \frac{f(k)}{f(-k)} = \frac{f(k)}{f^*(k^*)} \approx \frac{c}{c^*} \frac{k - k_R - i\Gamma}{k - k_R + i\Gamma} \approx e^{2i\delta_0} \frac{k - k_R - i\Gamma}{k - k_R + i\Gamma}$$

and we derive

$$\delta(k) = \delta_0 - \arctan \frac{\Gamma}{k - k_R} .$$

Thus we see that we are in presence of a resonance and that passing through the point k_R the phase-shift increases by π . It is worth noticing that being $\Gamma > 0$, in a neighborhood of k_R the phase shift is always increasing. To such a phase-shift there corresponds the well-known Breit-Wigner cross section which for $\delta_0 = 0$ is

$$\sigma = 4\pi \left| \frac{S - 1}{2ik} \right|^2 = \frac{4\pi}{k^2} \frac{\Gamma^2}{(k - k_R)^2 + \Gamma^2} .$$

It is interesting to examine the structure of the scattering wave function when the energy has a value near the real part of a resonance pole

$$\varphi(k, r) = \frac{1}{2ik} [f(k)f(-k, r) - f(-k)f(k, r)] .$$

Due to the Wronskian property, $\frac{\partial \varphi(k, r)}{\partial r}|_{r=0} = 1$ while for large r , $\varphi(k, r) \sim \frac{|f(k)|}{k} \sin(kr + \delta)$. For $k = k_R + i\Gamma$, $f(k) = 0$ while we always have $\frac{\partial \varphi(k, r)}{\partial r}|_{r=0} = 1$. Thus in proximity of a resonance i.e. $k = k_R$ we are in presence of a wave function which is large near the origin with an asymptotic normalization factor very small. The intuitive picture is that of a particle which interact for a long time with the scattering center.

13.4 Levinson's theorem

The following remarkable theorem holds: Given a potential $U(r)$ the difference between the phases of scattering at zero energy and at infinite energy equals $n\pi$ where n is the number of bound states. Such a theorem holds for all partial waves; here we prove it for the s -wave.

We know from $S(k) = f(k)/f(-k)$ that $S(0) = 1$ (unless we have a zero energy bound state in which case we have $S(0) = -1$ which we exclude) and also that $S(\infty) = 1$. Let us set by convention $\delta(\infty) = 0$ and define then δ by continuity. We know that $f(k)$ for $k \rightarrow \infty$ and $\text{Im } k \leq 0$ tends to 1. Then with this convention we have for real $k > 0$ $f(k) = \rho(k)e^{i\delta(k)}$ and $f(-k) = \rho(k)e^{-i\delta(k)}$ and at infinity in the lower half plane $\delta = 0$.

Let us consider now the function $\log f(k)$; we have that

$$\log f(+0) - \log f(-0) = \int_{-0}^{+0} \frac{f'(k)}{f(k)} dk$$

where the integral is performed along the contour represented in figure (13.4) enclosing all the zeros of $f(k)$. We normalize the logarithm so that $\log f(\infty) = 0$ and we have that

$$\log f(+0) - \log f(-0) = 2i\delta(0) .$$

The integrand is analytic except for the poles at the values of k corresponding to a zero of $f(k)$. These poles have residue 1 being the zeros of $f(k)$ in the lower

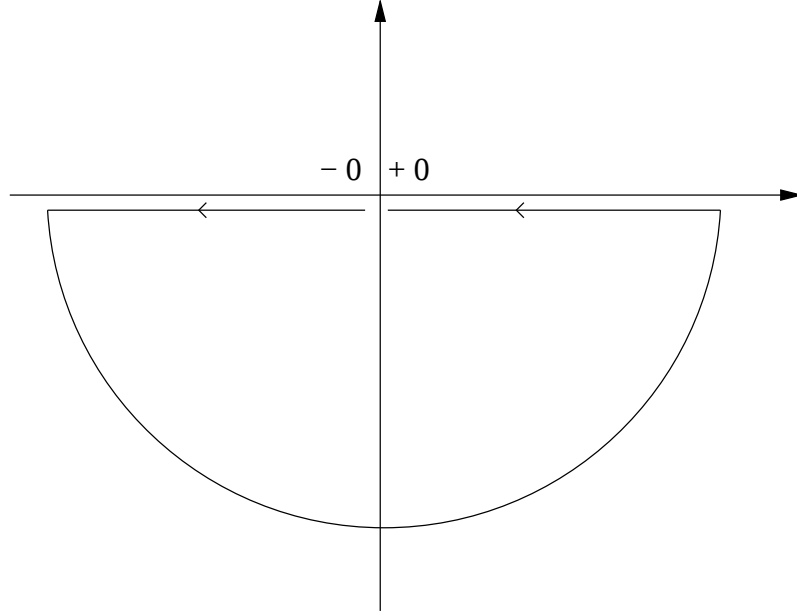


Figure 13.4: Integration contour

half-plane simple zeros (see section 13.2). As the zeros of $f(k)$ in the lower half plane represent the bound states of the system we have

$$\delta(0) = \pi n_B .$$

13.5 Construction of special potentials

It is instructive to examine potentials which give rise to particularly simple S -matrices. Let us try to construct an S -matrix of type

$$S = -\frac{k + i\chi}{k - i\chi} \quad (\chi \text{ real})$$

satisfying the condition $S(0) = 1$. This is obtained by imposing that the wave function be free except for the linear condition at the origin

$$\frac{\varphi'(0)}{\varphi(0)} = c . \tag{13.12}$$

Then from

$$\varphi(r) = e^{-ikr} - S(k)e^{ikr}$$

it follows

$$\frac{\varphi'(0)}{\varphi(0)} = -ik \frac{e^{-ikr} + S(k)e^{ikr}}{e^{-ikr} - S(k)e^{ikr}} \Big|_{r=0} = c .$$

Thus

$$S(k) = \frac{c + ik}{c - ik} .$$

which as we saw before corresponds for $c < 0$ to a bound state and for $c > 0$ to an antibound state.

A boundary condition of this type can be approximated by an attractive potential well of short range and large depth. One has simply to see whether, as the radius tends to zero, it is possible to change the depth in such a way as to obtain in the limit eq.(13.12).

Let us note that if the depth of the well is very large the wave number inside the well does not depend on the particle energy. We have for the wave number v inside the well, given k^2 and the potential depth $-U$

$$v^2 = k^2 - U$$

and

$$\frac{\varphi'(a)}{\varphi(a)} = \frac{v \cos va}{\sin va} .$$

If we choose

$$a = \frac{\pi}{2v} - \frac{c}{v^2}$$

we have for $-U \rightarrow +\infty$

$$\lim_{a \rightarrow 0} \frac{\varphi'(a)}{\varphi(a)} = \lim_{a \rightarrow 0} \frac{v \cos(\frac{\pi}{2} - \frac{c}{v})}{\sin(\frac{\pi}{2} - \frac{c}{v})} = c .$$

Thus the S -matrix so constructed is the limit case of a potential problem.

We want now to examine the problem of the construction of potentials which give rise to simple S -matrices in a systematic way [3].

Let us try to obtain Jost functions such that $g(k, r) = f(k, r)e^{ikr}$ is a simple algebraic function.

Setting $\varphi = e^{-ikr}\chi(k, r)$ from Schrödinger equation $-\varphi'' + U\varphi = k^2\varphi$, similarly to what done for Jost functions, we obtain the equation for $\chi(k, r)$

$$\chi'' - 2ik\chi' = U\chi .$$

Let χ be a solution of such equation with the property that there exists the limit of $\chi(k, r)$ for $r \rightarrow \infty$, and that such a limit is finite. We have

$$f(k, r) = e^{-ikr} \frac{\chi(k, r)}{\chi(k, \infty)}$$

and thus

$$f(k, 0) \equiv f(k) = \frac{\chi(k, 0)}{\chi(k, \infty)} .$$

The S -matrix is then given by

$$S(k) = \frac{f(k)}{f(-k)} = \frac{\chi(k, 0)}{\chi(-k, 0)} \frac{\chi(-k, \infty)}{\chi(k, \infty)} .$$

We start from the assumption that $\chi(k, r)$ be a polynomial in k .

0) Zero degree polynomial

$$\chi(k, r) = a(r) .$$

This gives rise to a Jost function which does not depend of k . From

$$a'' - 2ika' = Ua$$

it follows that, given the independence of k , $a' \equiv 0$ and thus $a'' \equiv 0$. Then $aU = 0$. If $U \neq 0$ we cannot construct the Jost function because we have $a = 0$. If $U = 0$ we have $\chi(k, r) = \text{const}$ and $f(k, r) = e^{-ikr}$ which is the Jost function of the free particle and $S(k) = 1$.

1) First degree polynomial

$$\chi(k, r) = b(r)k + a(r) .$$

From

$$\chi'' - 2ik\chi' = U\chi$$

it follows

$$kb'' + a'' - 2ik(b'k + a') = (kb + a)U .$$

In order to this to be true as k varies, it is necessary that $b' = 0$ i.e. $b = \text{const.}$

We can then write

$$\begin{aligned}\chi(k, r) &= 2k + ia(r) \\ ia'' + 2ka' &= (2k + ia)U\end{aligned}$$

and we have

$$U = a', \quad a'' = a'a . \quad (13.13)$$

The corresponding Jost function is

$$f(k, r) = e^{-ikr} \frac{2k + ia(r)}{2k + ia(\infty)}$$

while

$$[f(k, r)]^* = f(-k^*, r)$$

imposes

$$a^*(r) = a(r), \quad a^*(\infty) = a(\infty)$$

i.e. $a(r)$ has to be real. We must then look for a real solution of eq.(13.13). The first integration is immediate

$$a' - \frac{1}{2}a^2 = 2c . \quad (13.14)$$

To obtain a linear equation it is convenient to pose

$$w = \exp \left(-\frac{1}{2} \int_0^r a(r') dr' \right), \quad \text{with} \quad w > 0$$

and we have

$$a = -\frac{2w'}{w}, \quad a' = -2 \frac{w''w - w'^2}{w^2} .$$

In terms of w equation (13.14) becomes

$$w'' + cw = 0 .$$

In order for w to be real c has to be real. We recall furthermore that we are interested in solutions with $w(r) > 0$ for $r > 0$.

i) $c = 0$

$$w'' = 0, \quad w = \alpha_1 + \beta_1 r, \quad \alpha_1 > 0, \quad \beta_1 > 0$$

from which

$$\alpha_1 + \beta_1 r = \exp \left(-\frac{1}{2} \int_0^r a(r') dr' \right)$$

$$a(r) = -\frac{2}{\frac{\alpha_1}{\beta_1} + r} \equiv -\frac{2}{\alpha + r}$$

$$U(r) = a'(r) = \frac{2}{(\alpha + r)^2}.$$

The Jost function is given by

$$f(k) = \frac{k - i/\alpha}{k}$$

from which

$$S(k) = \frac{k - i/\alpha}{k + i/\alpha}.$$

Such an S -matrix has a single pole for $k = -i/\alpha$, which being $\alpha > 0$ is an antibound state. The fact that $f(k)$ has a pole for $k = 0$ is not surprising as a condition we used to prove $k = 0$ to be a point of analyticity of the Jost function $f(k)$ was

$$N(0) = \int_0^\infty x|U(x)|dx < \infty$$

a condition which is clearly violated by the potential in question.

ii) $c > 0$

We obtain an oscillating solution and thus w cannot be positive definite.

iii) $c < 0$

Let us write

$$c = -\lambda^2/4$$

and we have

$$w'' - \frac{\lambda^2}{4}w = 0, \quad w = e^{\frac{\lambda r}{2}} + \beta e^{\frac{-\lambda r}{2}}.$$

β has to be ≥ -1 otherwise there a point in which w is negative.

$$a = -\frac{2w'}{w} = \lambda \frac{\beta e^{-\lambda r} - 1}{\beta e^{-\lambda r} + 1}$$

from which

$$f(k) = \frac{k + i\frac{\lambda(\beta-1)}{2(\beta+1)}}{k - i\frac{\lambda}{2}} \equiv \frac{k + i\frac{\nu}{2}}{k - i\frac{\lambda}{2}} \quad (13.15)$$

where we set

$$\nu = \lambda \frac{\beta - 1}{\beta + 1}.$$

Notice that ν has the sign of $\beta - 1$ and $\nu < \lambda$ for $\beta \geq 1$.

$$U(r) = a' = -2\lambda^2\beta \frac{e^{-\lambda r}}{(1 + \beta e^{-\lambda r})^2} \quad (\text{Eckart potential}).$$

U is finite at the origin and at infinity behaves as $e^{-\lambda r}$.

The described treatment can be extended to $\chi(k, r)$ functions quadratic in k [3].

13.6 Discussion of the singularities of the S -matrix

The S -matrix in the case of Eckart potential is given by

$$S(k) = \frac{f(k)}{f(-k)} = \frac{k + i\nu/2}{k - i\nu/2} \frac{k + i\lambda/2}{k - i\lambda/2}.$$

We notice that the Jost function (13.15) has a pole for $k = i\lambda/2$. Thus we have that the Bargmann strip is of half-width $\lambda/2$ which is consistent with the asymptotic behavior $e^{-\lambda r}$ of the potential. Thus this pole of the S -matrix at $k = i\lambda/2$ is a typical interaction pole which lies at the boundary of the Bargmann strip. The Jost function has a zero at the point

$$k = -i\frac{\nu}{2} = -i\frac{\lambda}{2} \frac{\beta - 1}{\beta + 1}$$

to which there corresponds a pole of S at the point

$$k = i\chi = i\frac{\lambda}{2} \frac{\beta - 1}{\beta + 1}.$$

For $\beta > 1$ we have $\chi > 0$ and thus we have a bound state. If on the contrary $-1 < \beta < 1$, χ is negative and we have an antibound state. The residues of the S -matrix at the two poles $k = i\lambda/2$ and $k = i\nu/2$ are respectively

$$i\lambda \frac{\lambda + \nu}{\lambda - \nu} = i\lambda\beta$$

and

$$-i\nu \frac{\lambda + \nu}{\lambda - \nu} = -i\lambda \beta \frac{\beta - 1}{\beta + 1}$$

The phase shift is immediately obtained from

$$k \cot \delta = \operatorname{Re} \frac{2ik}{S(k) - 1} = -\frac{1}{a_0} + \frac{1}{2}r_0k^2 = -\frac{\nu\lambda}{2(\nu + \lambda)} + \frac{2}{\nu + \lambda}k^2. \quad (13.16)$$

Thus the phase-shift due to the Eckart potential obeys the effective range formula [4]. We can summarize the situation in the following table where the third and fourth column give the sign of $-i$ times the residues of the respective poles

β	ν	$i\lambda/2$	$i\nu/2$	a_0	r_0	state
$1 < \beta$	$0 < \nu < \lambda$	+	−	+	+	bound
$0 < \beta < 1$	$-\lambda < \nu < 0$	+	+	−	+	antibound
$-1 < \beta < 0$	$\nu < -\lambda < 0$	−	−	+	−	antibound

Table 13.1: Singularities of the S -matrix

Thus for $-1 < \beta < 0$ we have a repulsive potential which admits an antibound state which lies outside the Bargmann strip. For $0 < \beta < 1$ we have a weakly attractive potential with an antibound state within the Bargmann strip. For $\beta > 1$ we have strong attraction with the consequent formation of the bound state. The scattering length a_0 ($\delta \approx -a_0k$) is always positive except in the case $0 < \beta < 1$. Levinson's theorem is always satisfied as one sees from the table and eq.(13.16).

Of the four possibilities $a_0 < 0$, $a_0 > 0$, $r_0 < 0$, $r_0 > 0$ in the effective range formula all can be realized by the potential in question, except the $a_0 < 0$, $r_0 < 0$. To this there would correspond a resonant phase-shift with $\delta(\infty) - \delta(0) = \pi$.

References

- [1] R. Jost *Über die falschen Nullstellen der Eigenwerte der S-matrix*, Helvetica Physica Acta 20, 256 (1947)
- [2] T. Regge *Mathematical Theory of potential scattering*, Theoretical Physics IAEA (1963)

- [3] V. Bargmann *On the connection between phase shifts and scattering potentials*,
Rev. Mod. Phys. 21 488 (1949)
- [4] R.G. Newton *Scattering theory of waves and particles*, Dover Publications,
Mineola N.Y. (1982)

Chapter 14

The functional formulation of quantum mechanics

14.1 Introduction

Quantum mechanics in its final form was discovered in two different mathematical schemes which were soon proved to be equivalent. We can call it the operator formulation in so far as the dynamical variables are represented by linear operators which act on the linear space of states.

The quantization rule is provided by the correspondence principle i.e. replacing Poisson brackets of classical mechanics with commutators divided by $i\hbar$. We saw that this statement has to be taken with great care.

The hamiltonian formulation of classical mechanics is not explicitly relativistic invariant; in fact H , i.e. the energy, is not a relativistic invariant but the zero component of a four-vector.

On the contrary the classical Lagrangian approach is explicitly covariant insofar through the Lagrangian one succeeds to define an action which is an invariant

$$S = \int_{q_0 t_0}^{q t} L(q, \dot{q}) dt$$

i.e. the action has the same value in all reference frames. The vanishing of the variation of S , $\delta S = 0$, provides the equations of motion. Being S an invariant the equations of motion are covariant.

E.g. for a charged particle in an external electromagnetic field we have

$$S = \int_{q_0 t_0}^q \left(-mc \sqrt{\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} + \frac{e}{c} A_\mu \frac{dx^\mu}{d\lambda} \right) d\lambda$$

$$x^0 = ct, A_\mu = (\mathbf{A}, -cV).$$

The search for a lagrangian formulation of quantum mechanics and thus the possibility to give a relativistically invariant formulation of the same was the main motivation of Feynman [1]. It was inspired by some developments by Dirac [2].

The main characteristics of the functional formulation are

1. It is more intuitive, as far as quantum mechanics may be intuitive.
2. It requires mathematics of higher level: the concept of integration on a space of functions.
3. Only few problems can be solved directly in this approach.
4. It has a great heuristic and formal power:
 - 4.1 It gives a strong guide on how to organize the perturbative series.
 - 4.2 It possesses strong analogies with statistical mechanics.
 - 4.3 It provides a method to develop new approximation schemes.

14.2 The path integral

As in [1] we shall deal with non relativistic quantum mechanics. Instead of following the line historically developed by Feynman we shall follow the inverse path i.e. how from the usual operator hamiltonian formulation one can reach Feynman's one.

We shall refer to the one dimensional case as it is not difficult to go to higher dimensions. The object we take into consideration is the *propagator* which we already studied

$$G(q, t; q_0, t_0) = \langle q | e^{-\frac{i}{\hbar} H(t-t_0)} | q_0 \rangle$$

using Dirac's notations, where H is supposed to be time independent.

It is known that the Green function gives us the probability amplitude for a particle which at time t_0 is in q_0 to be found at time t in q ; the probability is

given by

$$|G(q, t; q_0, t_0)|^2.$$

The Green function allows more generally to compute transition amplitudes between states

$$\begin{aligned} \langle \phi | e^{-\frac{i}{\hbar} H(t-t_0)} | \psi \rangle &= \int \int \langle \phi | q \rangle dq \langle q | e^{-\frac{i}{\hbar} H(t-t_0)} | q_0 \rangle dq_0 \langle q_0 | \psi \rangle = \\ &= \int \phi^*(q) dq G(q, t; q_0, t_0) dq_0 \psi(q_0) \end{aligned}$$

and thus solves the dynamical problem of quantum mechanics.

We recall that for a free particle i.e. $H = \frac{p^2}{2m}$ we have for times $t > t_0$

$$G_0(q, t; q_0, t_0) = e^{-\frac{i\pi}{4}} \left(\frac{m}{2\pi\hbar(t-t_0)} \right)^{\frac{1}{2}} e^{\frac{i(q-q')^2 m}{2\hbar(t-t_0)}} \equiv \left(\frac{m}{2i\pi\hbar(t-t_0)} \right)^{\frac{1}{2}} e^{\frac{i(q-q')^2 m}{2\hbar(t-t_0)}}.$$

Notice moreover that inserting a complete set of eigenstates of the energy

$$G(q, t; q_0, t_0) = \sum_n \langle q | E_n \rangle \langle E_n | q_0 \rangle e^{-\frac{i}{\hbar} E_n(t-t_0)} + \int dE \langle q | E \rangle \langle E | q_0 \rangle e^{-\frac{i}{\hbar} E(t-t_0)}$$

and from Fourier analysis of $G(q, t; q_0, t_0)$ one succeeds in recovering the energy levels and the relative wave functions.

Notice that the same informations can be obtained from

$$G_E(q, \tau; q_0, 0) = \langle q | e^{-\frac{\tau}{\hbar} H} | q_0 \rangle = \sum_n e^{-\frac{\tau}{\hbar} E_n} \langle q | E_n \rangle \langle E_n | q_0 \rangle + \int dE e^{-\frac{\tau}{\hbar} E} \langle q | E \rangle \langle E | q_0 \rangle$$

and thus it is possible to recover the information about energy levels and states in terms of the anti-Laplace i.e. the Mellin transform.

Given a time t_1 with $t > t_1 > t_0$ we can also rewrite

$$G(q, t; q_0, t_0) = \langle q | e^{-\frac{i}{\hbar} H(t-t_1)} e^{-\frac{i}{\hbar} H(t_1-t_0)} | q_0 \rangle = \int G(q, t; q_1, t_1) dq_1 G(q_1, t_1; q_0, t_0). \quad (14.1)$$

Such an expression is reminiscent of the formula for probabilistic processes i.e.

$$P(q, t; q_0, t_0) = \int P(q, t; q_1, t_1) dq_1 P(q_1, t_1; q_0, t_0). \quad (14.2)$$

On the other hand in quantum mechanics

$$P(q, t; q_0, t_0) = |G(q, t; q_0, t_0)|^2$$

and thus if eq.(14.1) is true the (14.2) cannot be true. Actually (14.2) is true if at time t_1 we perform a measurement of the coordinate q and then leave again the system to evolve unperturbed from time t_1 to time t .

We can now insert many times, e.g. equally spaced, between t_0 and t and we have

$$G(q, t; q_0, t_0) = \int \cdots \int G(q, t; q_{n-1}, t_{n-1}) dq_{n-1} G(q_{n-1}, t_{n-1}; q_{n-2}, t_{n-2}) dq_{n-2} \cdots \\ \cdots G(q_2, t_2; q_1, t_1) dq_1 G(q_1, t_1; q_0, t_0)$$

with $t_k = t_0 + k\Delta$ and $\Delta = (t - t_0)/n \equiv \frac{T}{n}$. What did we gain with this decomposition? If the Hamiltonian is of form $H = V + K$ with $V = V(\hat{q})$ and $K = \hat{p}^2/2m$ we have

$$G(q_k, t_k; q_{k-1}, t_{k-1}) = \langle q_k | e^{-\frac{i}{\hbar}(V+K)\Delta} | q_{k-1} \rangle = \\ = \langle q_k | e^{-\frac{i}{\hbar}V\Delta} e^{-\frac{i}{\hbar}K\Delta} | q_{k-1} \rangle + o(\Delta)$$

where qualitatively $o(\Delta)/\Delta \rightarrow 0$ for $\Delta \rightarrow 0$.

This result can be understood from the Baker-Campbell-Hausdorff formula

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+O_3}.$$

Actually there exists a theorem known as Trotter formula which we shall prove in the next section which says that

$$e^{-\frac{i}{\hbar}HT} = s - \lim_{n \rightarrow \infty} (e^{-\frac{i}{\hbar}V\frac{T}{n}} e^{-\frac{i}{\hbar}K\frac{T}{n}})^n$$

where $s - \lim$ stays for strong limit. Being the error due to the separation of the two exponential of second order, in the limit in which n goes to infinity the error induced in the product goes to zero. In fact for the accumulated error in the multiplication of n terms we have $|(1 + \frac{\varepsilon}{n})^n - 1| \leq e^{|\varepsilon|} - 1$ which goes to zero for $\varepsilon \rightarrow 0$. The main difficulty in proving Trotter formula is in dealing with unbounded operators whose appearance is the rule in the Hamiltonians of quantum mechanics. We have

$$G(q, t; q_0, t_0) = \\ = \lim_{n \rightarrow \infty} \int \cdots \int \langle q | e^{-\frac{i}{\hbar}V\Delta} e^{-\frac{i}{\hbar}K\Delta} | q_{n-1} \rangle dq_{n-1} \langle q_{n-1} | e^{-\frac{i}{\hbar}V\Delta} e^{-\frac{i}{\hbar}K\Delta} | q_{n-2} \rangle dq_{n-2} \cdots \\ \cdots dq_1 \langle q_1 | e^{-\frac{i}{\hbar}V\Delta} e^{-\frac{i}{\hbar}K\Delta} | q_0 \rangle.$$

Let us consider the generic term

$$\begin{aligned} \langle q'' | e^{-\frac{i}{\hbar} V(\hat{q}) \Delta} e^{-\frac{i}{\hbar} \frac{\hat{p}^2}{2m} \Delta} | q' \rangle &= \int e^{-\frac{i}{\hbar} V(q'') \Delta} \langle q'' | e^{-\frac{i}{\hbar} \frac{\hat{p}^2}{2m} \Delta} | p' \rangle dp' \langle p' | q' \rangle = \\ &= \int e^{-\frac{i}{\hbar} V(q'') \Delta} e^{-\frac{i}{\hbar} \frac{p'^2}{2m} \Delta} \langle q'' | p' \rangle dp' \langle p' | q' \rangle = \left(\frac{m}{2\pi i \hbar \Delta} \right)^{\frac{1}{2}} e^{-\frac{i}{\hbar} V(q'') \Delta} e^{\frac{im(q''-q')^2}{2\hbar \Delta}} \end{aligned}$$

from which

$$\begin{aligned} G(q, t; q_0, t_0) &= \\ &= \lim_{n \rightarrow \infty} \left(\frac{m}{2\pi i \hbar \Delta} \right)^{\frac{n}{2}} \int e^{-\frac{i}{\hbar} V(q) \Delta} e^{\frac{im(q-q_{n-1})^2}{2\hbar \Delta}} dq_{n-1} e^{-\frac{i}{\hbar} V(q_{n-1}) \Delta} e^{\frac{im(q_{n-1}-q_{n-2})^2}{2\hbar \Delta}} dq_{n-2} \dots \\ &\quad dq_1 e^{-\frac{i}{\hbar} V(q_1) \Delta} e^{\frac{im(q_1-q_0)^2}{2\hbar \Delta}} . \end{aligned} \quad (14.3)$$

Given now the partition of the time interval $t-t_0$, with the times $t_n = t, t_{n-1}, \dots, t_1, t_0$ and taken arbitrary values of $q_{n-1}, \dots, q_1$ let us consider the motion given by the piecewise linear curve which at the time t_k equals q_k . We want to compute the classical action relative to such a motion. We have

$$\begin{aligned} \int_{q_0 t_0}^{q t} L(q, \dot{q}) dt &= \sum_{k=1}^n \left[- \int_{t_{k-1}}^{t_k} V \left(q_{k-1} + \frac{q_k - q_{k-1}}{\Delta} (t - t_{k-1}) \right) dt + \frac{m}{2} \left(\frac{q_k - q_{k-1}}{\Delta} \right)^2 \Delta \right] \approx \\ &\approx \sum_{k=1}^n \left[-V(q_k) \Delta + \frac{m}{2} \frac{(q_k - q_{k-1})^2}{\Delta} \right] \end{aligned}$$

where we performed an error of order of Δ^2 replacing

$$\int_{t_{k-1}}^{t_k} V(q(t)) dt \quad \text{with} \quad V(q_k) \Delta.$$

Such error is irrelevant in the limit of $n \rightarrow \infty$. The limit for $n \rightarrow \infty$ which we know to exist and which rigorously equals the Green function, is the definition of the functional integral on the paths which connect the event (q_0, t_0) to (q, t) . Thus we have reached the fundamental result

$$G(q, t; q_0, t_0) = \int \mathcal{D}[q(t)] e^{\frac{i}{\hbar} S[q(t)]}$$

where the symbol $\mathcal{D}[q(t)]$ stays for

$$\lim_{n \rightarrow \infty} \left(\frac{m}{2\pi i \hbar \Delta} \right)^{\frac{n}{2}} \prod dq_{n-1} \dots dq_1 .$$

Notice that L is not an operator but the classical Lagrangian of the system. Thus we do not have operators but functional integrals i.e. integrals on the space of functions.

All paths are possible and all equally probable, i.e. all contribute with the same weight which is simply a phase. The paths which are near the classical path, being on the classical path the action stationary give the most important contribution while those far away from the classical path tend to give contributions which cancel by interference.

In completely analogous way one can compute

$$G_E(q, \tau; q_0, 0) = \langle q | e^{-\frac{\tau}{\hbar} H} | q_0 \rangle = \lim_{n \rightarrow \infty} \left(\frac{m}{2\pi\hbar\Delta} \right)^{\frac{n}{2}} \int e^{-\frac{V(q)}{\hbar}\Delta} e^{-\frac{m(q-q_{n-1})^2}{2\hbar\Delta}} dq_{n-1} \dots \quad (14.4)$$

We see that

$$e^{-\frac{V(q_k)}{\hbar}\Delta} e^{-\frac{m(q_k-q_{k-1})^2}{2\hbar\Delta}} \approx e^{\frac{1}{\hbar} \int_{\tau_{k-1}}^{\tau_k} \left[-\frac{m}{2} (\dot{q}(\tau))^2 - V(q(\tau)) \right] d\tau} = e^{\frac{1}{\hbar} \int_{\tau_{k-1}}^{\tau_k} L_E(q, \dot{q}) d\tau}$$

where the “euclidean Lagrangian” $L_E(q, \dot{q})$ is defined by

$$L_E(q, \frac{\partial q}{\partial \tau}) = L(q, \frac{\partial q}{\partial(-i\tau)})$$

i.e. it is formally obtained by replacing t with $-i\tau$. This process is called the Wick rotation and intervenes in a variety of problems. The euclidean functional integral with periodic boundary conditions turns out to be a powerful calculational tool of the partition function of statistical mechanics of the system described by the Hamiltonian H as the partition function is given by

$$Z(\beta) = \text{Tr} (e^{-\beta H}) = \int dq G_E(q, \hbar\beta; q, 0) = \int dq \int \mathcal{D}[q(\tau)] e^{\frac{1}{\hbar} \int_{q,0}^{q,\hbar\beta} L_E(q, \dot{q}) d\tau}.$$

There exists also an interpretation of eq.(14.4) at the level of classical statistical mechanics: eq.(14.4) before taking the limit for $n \rightarrow \infty$ can be seen as the partition function of a one dimensional chain of classical degrees of freedom, like a spin chain, subject to the external potential $V(q)$ and first neighbors interaction with interaction energy $c(q_k - q_{k-1})^2$ and with periodic boundary conditions, being c a positive constant .

We follow now the inverse path i.e. we show how the Schrödinger equation is derived from the functional formulation of quantum mechanics. We shall use the following results which hold for positive α

$$\begin{aligned}\int e^{\frac{i\alpha q^2}{2}} dq &= e^{\frac{i\pi}{4}} \sqrt{\frac{2\pi}{\alpha}}, \quad \int e^{\frac{i\alpha q^2}{2}} q^{2n+1} dq = 0 \\ \int e^{\frac{i\alpha q^2}{2}} q^2 dq &= \frac{e^{\frac{3i\pi}{4}}}{\alpha} \sqrt{\frac{2\pi}{\alpha}}, \quad \int e^{\frac{i\alpha q^2}{2}} q^4 dq = -\frac{3e^{\frac{i\pi}{4}}}{\alpha^2} \sqrt{\frac{2\pi}{\alpha}}, \\ \int e^{\frac{i\alpha q^2}{2}} q^{2n} dq &= \frac{c(n)}{\alpha^n} \frac{1}{\sqrt{\alpha}}.\end{aligned}$$

We have from (14.3)

$$\begin{aligned}\psi(q, t + \Delta) &\approx e^{-\frac{i}{\hbar} V(q)\Delta} \times \left(\frac{m}{2\pi i \hbar \Delta}\right)^{1/2} \int e^{\frac{im(q-q')^2}{2\hbar \Delta}} dq' \psi(q', t) = \\ &\left(1 - \frac{i}{\hbar} V(q)\Delta + O(\Delta^2)\right) \times \\ &\left(\frac{m}{2\pi i \hbar \Delta}\right)^{1/2} \int e^{\frac{im(q-q')^2}{2\hbar \Delta}} dq' \left[\psi(q) + (q' - q)\psi'(q, t) + \frac{(q' - q)^2}{2} \psi''(q, t) + \dots \right] = \\ &\left(1 - \frac{i}{\hbar} V(q)\Delta + O(\Delta^2)\right) \left[\psi(q, t) + \frac{i\hbar \Delta}{2m} \psi''(q, t) + O(\Delta^2) \right] = \\ &\psi(q, t) - \frac{i}{\hbar} \left(-\frac{\hbar^2}{2m} \psi''(q, t) + V(q)\psi(q, t) \right) \Delta + O(\Delta^2).\end{aligned}$$

In this way we have recovered the Schrödinger evolution equation.

We add now a few remarks.

1. As already mentioned the same informations on the bound states and on the wave functions of the bound states can be obtained by studying the “euclidean” propagator $e^{-\frac{\tau}{\hbar} H}$ by computing the inverse Laplace transform i.e. the Mellin transform.

From the computational viewpoint the euclidean propagator is much simpler especially at the numerical level. This because the $\exp(\frac{1}{\hbar} \int L_E(q, \dot{q}) d\tau)$ has now the nature of a probability and thus one can apply to it all the results of statistical mechanics. Formally the transition from L to L_E is obtained setting $x^4 = ix^0 = i\tau$.

2. Configurations which are not minima but are stationary, can give particularly relevant contributions to the functional integral and thus can be taken as starting point of new methods of approximation.

14.3 Proof of Trotter formula

As we already mentioned the difficulty in proving Trotter formula is the occurrence in the Hamiltonian of unbounded operators.

Let A and B two self-adjoint operators on an separable Hilbert space $\mathcal{H}$ and let $A + B$ also be self adjoint, with domain $D(A) \cap D(B)$. A typical example is $A = \frac{\hat{\mathbf{p}}^2}{2m} = -\frac{\nabla^2}{2m}$ and $B = V(\mathbf{q})$ bounded; or $A = \frac{\hat{\mathbf{p}}^2}{2m} = -\frac{\nabla^2}{2m}$ and $B = V(\mathbf{q})$ bounded at infinity with a Coulomb singularity at the origin.

Theorem [6]:

$$e^{it(A+B)} = s - \lim_{n \rightarrow \infty} (e^{i\frac{t}{n}A} e^{i\frac{t}{n}B})^n \quad (14.5)$$

where $s - \lim$ stays for strong limit.

If A and B are lower bounded it holds also

$$e^{-t(A+B)} = s - \lim_{n \rightarrow \infty} (e^{-\frac{t}{n}A} e^{-\frac{t}{n}B})^n . \quad (14.6)$$

Proof: Let us define $S_t = e^{it(A+B)}$, $V_t = e^{itA}$, $W_t = e^{itB}$, $U_t = V_t W_t$. Also let us define the time evolved of ψ by $\psi_t = S_t \psi$.

We must prove that

$$\lim_{n \rightarrow \infty} \|(S_{\frac{t}{n}}^n - U_{\frac{t}{n}}^n) \psi\| = 0 \quad (14.7)$$

but as all intervening operators are of norm 1, it is sufficient to prove eq.(14.7) for all ϕ of a dense set and we shall choose $\phi \in D(A) \cap D(B)$.

We have

$$\|(S_{\frac{t}{n}}^n - U_{\frac{t}{n}}^n) \psi\| = \left\| \sum_{j=0}^{n-1} U_{\frac{t}{n}}^{(n-j-1)} (S_{\frac{t}{n}} - U_{\frac{t}{n}}) S_{\frac{t}{n}}^j \psi \right\| \quad (14.8)$$

e.g.

$$SSSS - UUUU = (S - U)SSS + U(S - U)SS + UU(S - U)S + UUU(S - U) .$$

From Stone theorem we have for a vector $\phi \in D(A) \cap D(B)$

$$\lim_{s \rightarrow 0} \frac{S_s - 1}{s} \phi = i(A + B)\phi$$

and

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{U_s - 1}{s} \phi &= \lim_{s \rightarrow 0} \left[\frac{V_s - 1}{s} \phi + \frac{W_s - 1}{s} \phi + (V_s - 1) \frac{W_s - 1}{s} \phi \right] \\ &= i(A + B)\phi \end{aligned}$$

because V_s , like W_s are strongly continuous group of unitary operators.

Thus on a fixed vector $\phi \in D(A) \cap D(B)$ we have with $\Delta_n \equiv n(S_{\frac{t}{n}} - U_{\frac{t}{n}})$

$$\lim_{n \rightarrow \infty} \Delta_n \phi = 0. \quad (14.9)$$

We can rewrite eq.(14.8) as

$$\begin{aligned} \|(S_{\frac{t}{n}}^n - U_{\frac{t}{n}}^n)\phi\| &= \left\| \sum_{j=0}^{n-1} U_{\frac{t}{n}}^{(n-j-1)} (S_{\frac{t}{n}} - U_{\frac{t}{n}}) \phi_{s_j} \right\| \leq \\ &\leq n \sup_{0 \leq s \leq t} \|(S_{\frac{t}{n}} - U_{\frac{t}{n}}) \phi_s\| = \sup_{0 \leq s \leq t} \|\Delta_n \phi_s\|. \end{aligned} \quad (14.10)$$

where $s_j = j \frac{t}{n}$.

If the vector ϕ_s were fixed, i.e. independent of s , the theorem would already be proven. The problem is that ϕ_s varies with s and an excessive non uniformity of the limit of eq.(14.10) could jeopardize the result i.e. it would be not assured that for $n \rightarrow \infty$ the last member of eq.(14.10) tends to zero.

One notices however that

1. $D \equiv D(A) \cap D(B)$, with the metric

$$|||\phi||| = \|\phi\| + \|(A + B)\phi\|$$

which is stricter than the one of the Hilbert space, has the structure of a Banach space. Such metric is sometime called the *graph metric*.

Note that D with the Hilbert metric $\|\cdot\|$ is not a Banach space as it is not complete (in general). Instead D with the metric $|||\cdot|||$ is complete, being $A + B$ a closed operator as a result of being $A + B$ self-adjoint.

2. Each Δ_n , considered as a linear operator from the Banach space D to $\mathcal{H}$ is such that for every $\phi \in D$ fixed

$$\sup_n |||\Delta_n \phi||| < \infty$$

and thus for the principle of uniform boundedness, i.e. the Banach-Steinhaus theorem, we have

$$\|\Delta_n \phi\| < C \|\phi\|, \quad \|(S_{\frac{t}{n}} - U_{\frac{t}{n}})\phi\| < \frac{C}{n} \|\phi\|,$$

Given $\phi \in D$, for s varying in $[0, t]$ ϕ_s varies continuously in D in the metric $\|\cdot\|$ because first

$$(A + B)e^{i(A+B)s}\phi = e^{i(A+B)s}(A + B)\phi$$

and thus $\phi_s \in D$ and then for the strong continuity of the unitary group $e^{i(A+B)s}$. As a consequence given $\varepsilon > 0$ we can divide $[0, t]$ in a finite number of intervals $[0, t_1), [t_1, t_2) \dots [t_{N-1}, t]$ such that within each interval ϕ_s can be written as the sum of a vector which does not depend on s and a vector ζ_s whose $\|\cdot\|$ norm is less than ε for all s . We have from (14.9) that the contribution of the constant components go to zero for $n \rightarrow \infty$.

For the contribution of the component ζ_s we have

$$\left\| \sum_{j=0}^{n-1} U_{\frac{t}{n}}^{(n-j-1)} (S_{\frac{t}{n}} - U_{\frac{t}{n}}) \zeta_{s_j} \right\| \leq \left\| \sum_{j=0}^{n-1} (S_{\frac{t}{n}} - U_{\frac{t}{n}}) \zeta_{s_j} \right\| \leq C \max \|\zeta_{s_j}\| < C\varepsilon$$

and due to the freedom in the choice of ε we have proved eq.(14.5).

The same proof holds also for the euclidean version (14.6), provided the potential $V(\mathbf{q})$ is lower bounded.

References

- [1] R.P. Feynman *Space-Time Approach to Non-Relativistic Quantum mechanics*, Rev. Mod. Phys. **20** 367 (1948)
- [2] P.A.M. Dirac *The principles of quantum mechanics*, Chap.4, par.32, Clarendon Press, Oxford (1958)
- [3] W. Pauli *Lectures on theoretical physics*, The Massachusetts Institute of Technology, Boringhieri (1973)
- [4] T.D.Lee *Particle physics and introduction to field theory*, Harwood Academic Publishers, New York

- [5] S. Weinberg *The quantum theory of fields*, Vol.I Cambridge University Press (1995)
- [6] B. Simon *Functional integration and quantum physics*, Academic Press, New York, (1979)

Chapter 15

The violation of Bell's inequalities

15.1 Introduction

Let us consider a pair of non identical particles of spin $1/2$ which are produced in the singlet state and thus with state vector

$$|s\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_2 |+\rangle_1) \psi(q_1, q_2, t)$$

where

$$\psi(q_1, q_2, t) = \psi_1(q_1, t) \psi_2(q_2, t)$$

represents two wave packets which move apart in opposite directions.

After the wave packets have separated to large distance we perform the measure of $s_z^{(1)}$. We know that the possible results are $+1/2$ and $-1/2$ which in the following we shall denote as $+$ and $-$. The probability to obtain such results are $\langle s | P_1^+ | s \rangle$ and $\langle s | P_1^- | s \rangle$ where $P_1^+ = |+\rangle_1 \langle +|_1$ and $P_1^- = |-\rangle_1 \langle -|_1$.

According to the theory of measure if we obtain as a result $+$ the state of the system after the measure is represented by the vector $P_1^+ | s \rangle = \frac{1}{\sqrt{2}} |+\rangle_1 |-\rangle_2 \psi(q_1, q_2, t)$ which in order to perform further calculations is to be normalized to 1.

If we perform now a measure of $s_z^{(2)}$ we obtain with certainty, i.e. with probability 1 the value $-$. As when we performed the measure of $s_z^{(1)}$ on particle 1 such a

particle i.e. counter (1), was very far from particle 2, it is reasonable to maintain that such measuring process does not influence the system “particle 2”.

In different words as during the process of measure of 1 we have not disturbed particle 2, one can maintain that particle 2 had the value — for $s_z^{(2)}$ even before the measure of the spin of particle 1.

This circumstance led Einstein, Podolsky, Rosen (EPR [1]) to formulate the following criterion of physical reality:

If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there is an element of physical reality corresponding to this physical quantity.

From such a criterion we would conclude that the description of the state by means of the vector $|s\rangle$ is not complete because the value — for $s_z^{(2)}$ of particle 2, even if it has a physical reality, cannot be forecasted from the structure of the state $|s\rangle$.

EPR adopt the following criterion of completeness of a physical theory: *Every element of the physical reality must have a counterpart in the physical theory.*

From such a criterion it would appear that the state vector gives an incomplete description of the system, a description of statistical character as it happens e.g. in classical statistical mechanics where the density function $\rho(q, p)$ in phase space gives an incomplete description of a classical physical system.

Such considerations may put in doubt the completeness of quantum mechanics as a theory; they do not put in doubt the correctness of the predictions, in general of statistical character, of quantum mechanics.

However J.S. Bell in 1964 [2] has shown that the predictions of quantum mechanics violate in quantitative manner, certain relations imposed by the idea of reality given by EPR, which is a refinement of the usual idea of reality as something which exist independently of the observer, coupled with the requirement of locality.

Thus while EPR in their paper leave open the possibility of a completion of quantum mechanics in the sense of a realistic local theory Bell's analysis denies such a possibility.

The experiments confirm the predictions of quantum mechanics.

15.2 Bell's original argument

Let us consider two non identical particles of spin $1/2$ created in a singlet state

$$|s\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2).$$

We are interested in the mean value of the operator $\boldsymbol{\sigma}_1 \cdot \mathbf{a} \ \boldsymbol{\sigma}_2 \cdot \mathbf{b}$. Such a mean value is easily computed to be

$$P(\mathbf{a}, \mathbf{b}) \equiv \langle s | \boldsymbol{\sigma}_1 \cdot \mathbf{a} \ \boldsymbol{\sigma}_2 \cdot \mathbf{b} | s \rangle = -\mathbf{a} \cdot \mathbf{b} \quad (15.1)$$

This result is experimentally confirmed. We notice that according to quantum mechanics whenever we measure the spin of particle 1 along a direction $\mathbf{c}$ the result is $+\frac{1}{2}$ or $-\frac{1}{2}$, and if we measure then the spin of particle 2 along the same direction the result is $-\frac{1}{2}$ or $+\frac{1}{2}$ with strict correlation. As such strict result is predicted for arbitrary space separations even outside the light-cone one is tempted to provide a local description of the phenomenon i.e. that at the production the two particle receive a, in general unpredictable, imprinting λ and that the results of later measurements are the outcome of such an imprinting λ . Let $A(\mathbf{a}, \lambda)$ be result of the measure of $\boldsymbol{\sigma}_1 \cdot \mathbf{a}$ and similarly $B(\mathbf{b}, \lambda)$ the result of the measure of $\boldsymbol{\sigma}_2 \cdot \mathbf{b}$. We have $A(\mathbf{a}, \lambda) = \pm 1$ and $B(\mathbf{b}, \lambda) = \pm 1$. In this description we have for the mean value (15.1)

$$P(\mathbf{a}, \mathbf{b}) = \int d\lambda \rho(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda)$$

where the integral of ρ , which is a non negative function, is normalized to 1. As we know from quantum mechanics that $P(\mathbf{a}, \mathbf{a}) = -1$ due to the fact that A and B are in modulus bounded by 1 we have that $B(\mathbf{a}, \lambda) = -A(\mathbf{a}, \lambda)$. We can compute now the mean value (15.1) along an other direction $\mathbf{c}$ and subtracting obtain

$$\begin{aligned} P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) &= - \int d\lambda \rho(\lambda) A(\mathbf{a}, \lambda) (A(\mathbf{b}, \lambda) - A(\mathbf{c}, \lambda)) = \\ &= - \int d\lambda \rho(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) (1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)) \end{aligned}$$

having used $A(\mathbf{b}, \lambda)^2 = 1$. Thus

$$\begin{aligned} |P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| &\leq \int d\lambda \rho(\lambda) |A(\mathbf{a}, \lambda)(A(\mathbf{b}, \lambda)| \left(1 - A(\mathbf{b}, \lambda)A(\mathbf{c}, \lambda)\right) \\ &= 1 + \int d\lambda \rho(\lambda) A(\mathbf{b}, \lambda)B(\mathbf{c}, \lambda) = 1 + P(\mathbf{b}, \mathbf{c}) . \end{aligned}$$

As a result we have

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq 1 + P(\mathbf{b}, \mathbf{c}) . \quad (15.2)$$

Let us consider now the case in which $\mathbf{a} = (0, 0, 1)$, $\mathbf{b} = (0, 1, 0)$, $\mathbf{c} = (0, c_2, c_3)$ for which we obtain

$$\sqrt{1 - c_2^2} \leq 1 - c_2$$

which is violated for $c_2 > 0$. Thus we have a definitive disagreement between quantum mechanics and the prediction of a general local theory.

15.3 The Clauser-Horne-Shimony-Holt inequality

The inequalities of type (15.2) can be generalized as follows: The preparing apparatus produces a pair of systems with characteristics described by a parameter λ which will change for every emission with a probability distribution $f(\lambda)$.

It will be useful to consider the following product of difference of probabilities

$$(P_1(\text{yes}, a, \lambda) - P_1(\text{no}, a, \lambda)) \times (P_2(\text{yes}, b, \lambda) - P_2(\text{no}, b, \lambda))$$

where $P_1(\text{yes}, a, \lambda)$ is the probability to find the answer “yes” when the system 1 emitted by the producing apparatus enters the counter 1 set on the characteristic a (e.g. orientation) and so on. In principle the relation between λ and the result of the measurement may be also probabilistic.

It is useful to write

$$A_1(a, \lambda) = P_1(\text{yes}, a, \lambda) - P_1(\text{no}, a, \lambda)$$

$$A_2(b, \lambda) = P_2(\text{yes}, b, \lambda) - P_2(\text{no}, b, \lambda)$$

and as probabilities are positive numbers between 0 and 1 we have

$$|A_1(a, \lambda)| \leq 1, \quad |A_2(b, \lambda)| \leq 1 .$$

Consider now the mean value of the product of the above quantities i.e.

$$\begin{aligned}
 E(a, b) &= \int d\lambda f(\lambda) A_1(a, \lambda) A_2(b, \lambda) \\
 &= \int d\lambda f(\lambda) (P_1(\text{yes}, a, \lambda) - P_1(\text{no}, a, \lambda)) (P_2(\text{yes}, b, \lambda) - P_2(\text{no}, b, \lambda)) \\
 &= P(\text{yes}, a, \text{yes}, b) + P(\text{no}, a, \text{no}, b) - P(\text{yes}, a, \text{no}, b) - P(\text{no}, a, \text{yes}, b) .
 \end{aligned} \tag{15.3}$$

We have

$$E(a, b) + E(a, b') = \int d\lambda f(\lambda) A_1(a, \lambda) (A_2(b, \lambda) + A_2(b', \lambda))$$

and thus

$$|E(a, b) + E(a, b')| \leq \int d\lambda f(\lambda) |A_2(b, \lambda) + A_2(b', \lambda)|$$

and changing the setting of the counters and taking the difference instead of the sum

$$|E(a', b) - E(a', b')| \leq \int d\lambda f(\lambda) |A_2(b, \lambda) - A_2(b', \lambda)|$$

and adding

$$\begin{aligned}
 &|E(a, b) + E(a, b')| + |E(a', b) - E(a', b')| \leq \\
 &\int d\lambda f(\lambda) \left(|A_2(b, \lambda) + A_2(b', \lambda)| + |A_2(b, \lambda) - A_2(b', \lambda)| \right) \leq 2
 \end{aligned} \tag{15.4}$$

because

$$|A_2(b, \lambda) + A_2(b', \lambda)| + |A_2(b, \lambda) - A_2(b', \lambda)| \leq 2.$$

Eq. (15.4) is the Clauser-Horne-Shimony-Holt inequality [4]. It holds for any pair of produced systems 1 and 2 and for whatever type of counters 1 and 2. Notice that such inequality holds even if the fundamental theory were of statistical nature but always local. In fact the result of the measure of system 1 by counter 1 will still be described by a probability function $P_1(x, a, \lambda)$. Thus system 1 carries with himself through the value of λ , the imprinting of how even stochastically, it has been produced but the result of the measure by the counter 1 will not necessarily be well determined.

We now apply such inequality to the case we have been studying in the previous section. Let us consider again the state

$$|s\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) .$$

and specify the experiment as follows: The produced system is the one described at the beginning of section (15.1), the two counters are Stern-Gerlach polarimeters. x is the axis along which the two counters and the preparation device are located. The parameters a and b of the counters 1 and 2 are the rotation angles α and β around the x axis, of the two polarimeters with respect to a reference direction z . “yes” means that $+$ has been measured, “no” means that $-$ has been measured.

As $|s\rangle$ is invariant under rotations as far as the spin state vector is concerned, $P(\cdot, a; \cdot, b)$ of (15.3) can depend only on the difference of the two angles β and α . The quantum mechanical computation gives

$$P(\text{yes } a; \text{yes } b) = |{}_1\langle + | {}_2\langle +, \beta - \alpha | s \rangle|^2$$

where

$$|+, \beta - \alpha\rangle_2 = \left(\cos \frac{\beta - \alpha}{2} - i\sigma_{(2)}^x \sin \frac{\beta - \alpha}{2} \right) |+\rangle_2$$

i.e.

$$P(\text{yes } a; \text{yes } b) = \frac{1}{2} \sin^2\left(\frac{\beta - \alpha}{2}\right) = P(\text{no } a; \text{no } b)$$

and similarly we find

$$P(\text{yes } a; \text{no } b) = \frac{1}{2} \cos^2\left(\frac{\beta - \alpha}{2}\right) = P(\text{no } a; \text{yes } b) = \frac{1}{2} \cos^2\left(\frac{\beta - \alpha}{2}\right).$$

From these we can compute the $E(\cdot, \cdot)$.

$$E(a, b) = -\cos(\beta - \alpha).$$

We shall choose $\alpha = 0$, $\alpha' = \pi/2$, $\beta = \pi/4$, $\beta' = -\pi/4$ i.e. the individual counter can assume two positions one orthogonal to the other; the position of the counter 2 are rotated of $\pi/4$ with respect to those of counter 1.

We see that the combination

$$|E(a, b) + E(a, b')| + |E(a', b) - E(a', b')| = 2 \cos \frac{\pi}{4} + 2 \cos \frac{\pi}{4} = 2\sqrt{2} > 2 \quad (15.5)$$

violates the CHSH inequality (15.4). This shows that quantum mechanics predicts correlations which are higher than those we can expect from any realistic local theory. If one relaxes the principle of locality it is easy to show with examples that the CHSH inequality can be violated and actually one can supply

examples in which the violation of such inequalities is even stronger than that of eq.(15.5).

The results of quantum mechanics, confirmed by experiments, teach us that, as maintained by Bohr, the two particles 1 and 2 even when the wave packets $\psi_1(q_1, t)$ and $\psi_2(q_2, t)$ are far apart, cannot be considered as two independent systems but have to be considered as a unique quantum system; this is true until we perform an operation of measure.

These experiments can be performed even “outside the light-cone” i.e. in such a way as to exclude, according the principle of relativity, a transfer of information from counter 1 and counter 2 which would invalidate the proof of CHSH inequality [5].

Bell has shown, in the quantum field theory framework, that such correlations, higher than those which we can admit in any local realistic theory, do not allow the transmission of information with a speed higher than that of light i.e. the impossibility of constructing a superluminal telegraph by exploiting the violation of the CHSH inequalities. At the basis of such impossibility lies the statistical nature of the relationship between the state vector and the results of an experiment.

Even more stringent inequalities, which are violated by the predictions of quantum mechanics, can be deduced by studying triple correlations [6].

Refeces

- [1] A. Einstein, B. Podolsky, N. Rosen *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. 47, 777 (1935).
- [2] J.S. Bell *Speakable and unspeakable in quantum mechanics*, Cambridge University Press. (1987).
- [3] B. d’Espagnat *The quantum theory and reality*, Scientific American p.128 November 1979.
- [4] J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. 23 880 (1969), Erratum Phys. Rev. Lett. 24, 549 (1970).

- [5] A. Aspect, J. Dalibard, and G. Roger *Experimental test of Bell's inequalities using time-varying analyzers*, Phys. Rev. Lett. 49, 1804 (1982).
- [6] N.D. Mermin *Quantum mysteries revisited*, Am.J.Phys. 58, 731 (1990).